Blind Rectification of Radial Distortion by Line Straightness

Burak Benligiray, Cihan Topal
Anadolu University, Department of Electrical and Electronics Engineering, Eskisehir, Turkey
{burakbenligiray, cihant}@anadolu.edu.tr

***Abstract*—Lens distortion self-calibration estimates the distortion model using arbitrary images captured by a camera. The estimated model is then used to rectify images taken with the same camera. These methods generally use the fact that built environments are line dominated and these lines correspond to lines on the image when distortion is not present. The proposed method starts by detecting groups of lines whose real world correspondences are likely to be collinear. These line groups are rectified, then a novel error function is calculated to estimate the amount of remaining distortion. These steps are repeated iteratively until suitable distortion parameters are found. A feature selection method is used to eliminate the line groups that are not collinear in the real world. The method is demonstrated to successfully rectify real images of cluttered scenes in a fully automatic manner.**

***Index Terms*—Camera calibration; radial distortion; plumb-line method; distortion rectification; self-calibration.**

I. Introduction

Calibrating the camera under controlled conditions, e.g., using a calibration pattern, is called pre-calibration, and is well studied [1], [2], [3]. Such methods generally estimate the lens distortion along with the camera model and provide accurate results. However, pre-calibration may not viable in all situations. For example, Collins and Tsin argue that the intrinsic parameters of the camera may be affected due to handling or temperature and humidity conditions, thus the camera should be calibrated on-site [4]. Furthermore, one may not be able to use known calibration objects in the scene for calibration, or the camera may be unavailable for capturing special calibration images (a common case for images collected from the Internet). Using a set of arbitrary images for camera model estimation is called self-calibration [5].

Self-calibration generally uses multiple views to estimate the camera model [5], [6], [7]. A further specification of the application limitations is using a single image. A large set of studies estimate the lens distortion with a single image, using the projective invariant which states that lines in the real world are projected to lines in the image plane [8], [9], [10], [11], [12]. This approach is commonly known as the plumb-line method [13]. A similar assumption is that projections of parallel lines in real world intersect at a single vanishing point on the image [14], [15]. Plumb-line methods require real lines to be present on the scene and the vanishing point based methods require some of these lines to be parallel and not collinear. Due to their dependencies on the scene composition, it is difficult to make a fair comparison among these methods.

Plumb line methods use edges [8], [9], [10], [16], [17], [18] or lines [11], [14], [19] as lowest level features. The collection of these features are either modelled as a line, considering they would correspond to a line if the distortion was corrected [8], [9], [11], [17], [18], or they are modelled as a circular arc, considering a line distorted by a first degree lens distortion function is a circular arc [10], [16], [19]. Ideally, these low level features should belong to lines in the real world. Fully automatic methods tend to use a robust feature selection method such as RANSAC [11], [14], [16], [19] or LMedS [8]. Other methods either explicitly state that human intervention is necessary for feature selection, or since they are using a non-linear optimization algorithm for distortion parameter estimation, having a majority of valid features allows for convergence to an adequate solution.

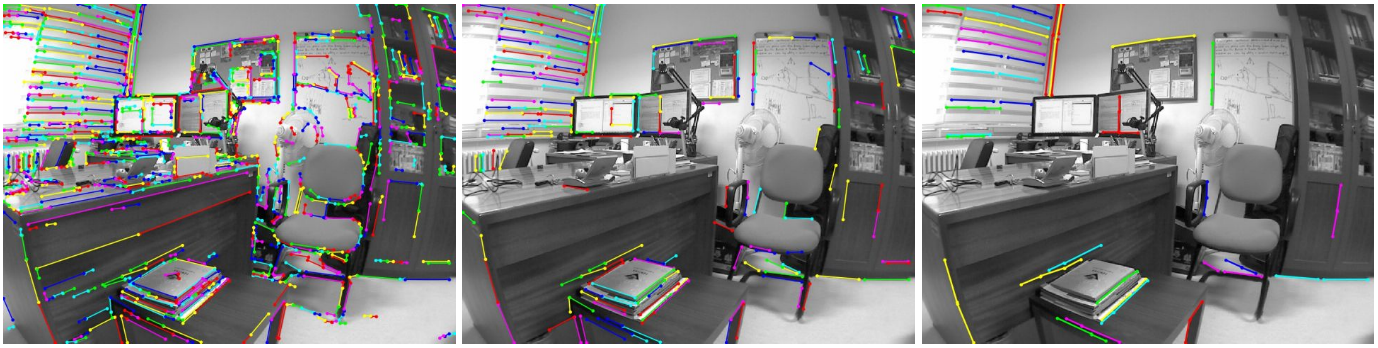
Estimation of parameters can be done by a closed form solution [8], [19], a linear solution [19], exhaustively [20], but the least squares approach is the most common [9], [10], [11], [15], [16], [18], [19], [21]. An error function to be minimized needs to be defined for using a least squares approach. Majority of the studies use the distances between the features and fitted models as their error function. Strand and Hayman argue that minimizing distances in the undistorted image causes a bias towards downscaling distortion parameters where all distances are reduced [19]. A zoom factor that scales the undistorted image is used by Alvarez et al. that aims to reduce this effect [17].

II. Estimation of Distortion Parameters

In this study, we will use the polynomial distortion model with two radial parameters, as done in [11], [12], [14], [15], [17], [21]. Let us assume $(\tilde{x}_d, \tilde{y}_d)$ are distorted screen coordinates and $(\tilde{x}_u, \tilde{y}_u)$ are rectified screen coordinates. These coordinates can be normalized by subtracting the distortion center, (c_x, c_y), thus yielding (x_d, y_d) and (x_u, y_u), respectively. The distances of the normalized points from distortion centers give the radii, r_d and r_u. According to this notation, the distortion that maps rectified coordinates to distorted coordinates is modelled as:

$$r_d = r_u(1 + \kappa_1 r_u^2 + \kappa_2 r_u^4) \tag{1}$$

$$\begin{aligned} x_d &= x_u(1 + \kappa_1 r_u^2 + \kappa_2 r_u^4) \\ y_d &= y_u(1 + \kappa_1 r_u^2 + \kappa_2 r_u^4) \end{aligned} \tag{2}$$



(a) All line segments. (b) Remaining line segments after elimination. (c) Grouped line segments.

Fig. 1: Steps of the feature extraction step. Endpoints of line segments are indicated by circular markers. Some of the resulting line groups are not collinear in real world. This issue is later addressed using a feature selection method.

The proposed method starts with line segment detection in the distorted image. These line segments are eliminated based on heuristics that will be detailed in Section II-A and the remaining sufficiently collinear ones are clustered into line groups. It is assumed that the environment is line dominant, hence the majority of the line groups belong to real world lines. The line groups that contradict this assumption will be eliminated later on. The line groups are rectified using a set of distortion parameters as described in Section II-B. The relative orientations of rectified line groups are assessed using the error function discussed in Section II-C. The distortion parameters are adjusted iteratively, until the error function converges to a minimum. Section II-D explains the method of eliminating the line groups that do not belong to real world lines. Assume there are N line groups and a minimum error function result is obtained following the prior steps. The same steps are applied for the $N-1$ combinations of the line groups. If all of the line groups belong to real world lines and are correctly localized, it is expected that N line groups yield the minimum error function value. Then, there are no line groups to be eliminated and the method is finished. If one of the $N-1$ combinations of the line groups have yielded the minimum error function result, the line group which is absent in this combination is eliminated, as it is inconsistent with the rest of the line groups. This line group removal operation is repeated until removing any more line groups does not improve the result.

A. *Feature Extraction*

The proposed method starts with line segment detection (see Fig. 1.a) [22]. Short features are reported to deteriorate performance in the literature [9], [11], [15], [16], [21], which is consistent with our experience, hence they are eliminated. Assuming the distortion center is close to the center of the image, line segments directed towards the center of the image are eliminated, because they will not contribute meaningful information regarding radial distortion (see Fig. 1.b).

The line segment detection method starts by detecting edge segments, then detecting line segments on these edge segments [22]. We can group the line segments that are extracted from the same edge segment. This method gives more reliable results, compared to distance and relative orientation based heuristics similar to the ones used in [11]. The line segments which are not grouped with at least one other line segment are also eliminated. Majority of the resulting chains of line segments are expected to belong to lines in real world (see Fig. 1.c). The elimination of false line groups will be addressed to in Section II-D.

B. *Iterative Rectification*

Rectifying the whole image is an expensive operation. Instead of rectifying the image at every parameter estimation iteration step, rectifying only the extracted features will be preferable. By doing so, we will obtain the distortion model that maps distorted points to rectified points. Note that the distortion model that is used for image rectification (see Eq. 1,2) is the inverse of this mapping function. When rectifying an image, our aim is to fill a set of blank pixels on the rectified image plane. Therefore, the function that maps the rectified points to distorted points must be found.

In the literature, the inverse distortion model is used throughout the self-calibration method, then the forward model is approximated using the inverse model as a final step [12], [14], [17]. This approach introduces a considerable amount of approximation error to the system. There is also an alternative distortion rectification method that directly uses the inverse parameters and gives better results [23]. However, this method cannot provide forward model parameters, which may be needed in some cases. As an alternative to these methods, we search for forward model parameters through the iterations. To rectify the line groups, we approximate the inverse distortion model at each step using Newton-Raphson method. By doing so, the approximation error caused by Newton-Raphson method is spread among the distortion parameter estimation iterations, hence the distortion parameters are chosen accordingly to compensate for some of this error. Since the result of the proposed method is forward distortion model parameters, they can be used for distortion rectification without needing to approximate their inverses.

While applying rectification, the distortion center is fixed to image center and the ratio of horizontal and vertical focal lengths is assumed to be unity. These parameters can easily be included to the proposed method as additional distortion model parameters. However, these factors are already close to the assumed values. Reaching to a better estimation requires a large number of well localized features, which are not present in our case with single images.

C. Error Function

An error function is used to measure the success of a rectification. Hence, only rectifying the image with more suitable distortion parameters must decrease the result of this function. As mentioned in Section I, distances from the fitted line or circle models are used to calculate the error function in the literature. This is not preferable, as the total zooming in or out effect of the applied rectification causes a bias in the error function. A rectification that causes the image to become smaller will also decrease all distances in the image. If the error function is directly related to the magnitude of model fitting error distances, it will have a positive bias towards distortion parameters that cause downscaling rectifications.

As a result of the feature extraction and clustering scheme discussed in Section II-A, we have multiple chains of line segments detected on the distorted image. Using the method described in Section II-B, these line segments are rectified. In the ideal case, all line segments of a chain will lie on a single line. This is commonly tested by fitting a line (or a circle) to these line segments and accumulating the distances from this model. In our method, we accumulate the relative orientations (see Fig. 2) of line segments in every chain (see Eq. 3), where θ is the relative orientation). This is more advantageous than the distance based error function, as it is not prone to the zoom-related bias discussed earlier. The mean square error is calculated as shown in Eq. 4 to find the error of a line segment chain.

$$\theta_i = arccos\left(\frac{\vec{l_i} \cdot \vec{l_{i+1}}}{\|\vec{l_i}\| \|\vec{l_{i+1}}\|}\right) \tag{3}$$

$$\varepsilon = \frac{1}{N}\sum_{i=1}^{N}\theta_i^2 \tag{4}$$

The function in Eq. 4 is only defined for a single line segment chain. There are multiple chains in an image. We do not weigh the effect of these chains based on the number of line segments or their total lengths, as they do not appear to have a significant meaning. Therefore, the total error function for an image is the summation of error functions for each individual line segment chain. An error function value for multiple images can be calculated in a similar manner to work with a larger number of features. In this study, we assumed that only one image is available.

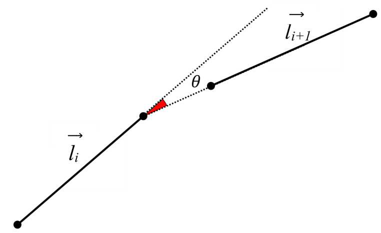


Fig. 2: θ, the relative orientation of two consecutive line segments.

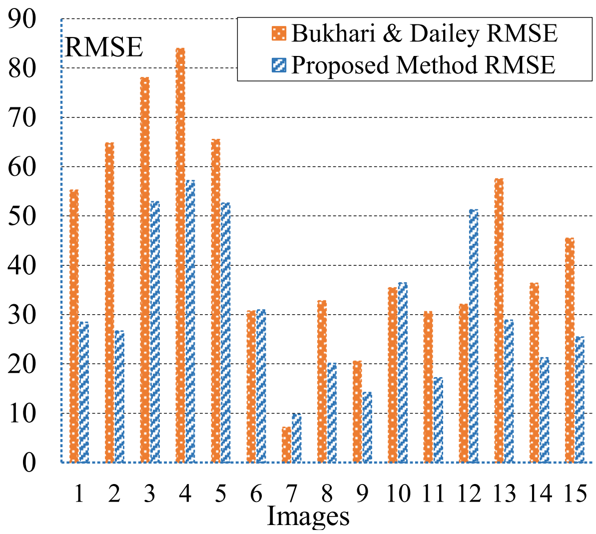
D. Feature Selection and Distortion Estimation

The error function presented in Section II-C is minimized using a non-linear optimization scheme [24]. At this point, the result will be acceptable if lines were dominant in the scene. Eliminating line groups that come from non-collinear elements will further improve the results. To eliminate the remaining line groups, sequential backward selection is applied. At first, all N line groups are used to estimate the distortion. Then, the distortion is estimated again by using the $N-1$ combinations of these line groups. The process stops if removal of any line group does not yield a better result than not removing any line group. Otherwise, the line group whose absence yields the least error function result is eliminated. This process is repeated until no further improvement can be done by removing line groups. For our test images, this method needs to repeat the non-linear optimization $\sim$ 30 times for each frame. As expected, there is a tradeoff between robustness and running time. Alternatively, RANSAC is commonly used in the literature for robust selection of features [11], [14], [16], [19]. While also assuming that line features coming from real world lines are high in number, RANSAC may fail to come up with a correct selection of features, regardless of its number of maximum iterations. The proposed feature selection method has the advantage of being deterministic and providing more reliable results.

III. Experimental Results

The experiments are done using 15 real images taken with a camera with significant barrel distortion. The camera is calibrated with a checkerboard pattern using Bouguet's calibration toolbox [25] and the test images are rectified using these results. The rectified images are considered to be the ground truths, as pre-calibration methods tend to give fairly accurate results. This step included estimating tangential distortion and distortion center to obtain the best result possible.

We used Bukhari and Dailey's automatic radial lens distortion rectification implementation for comparison [16]. This method estimates a single distortion parameter and the distortion center. Since the pre-calibration method also estimates the distortion center, the difference will cause a disadvantage against our results. RANSAC causes Bukhari and Dailey's method to be non-deterministic. The algorithm sometimes gives invalid results, such as largely empty images. We repeated the procedure until an acceptable result was achieved.



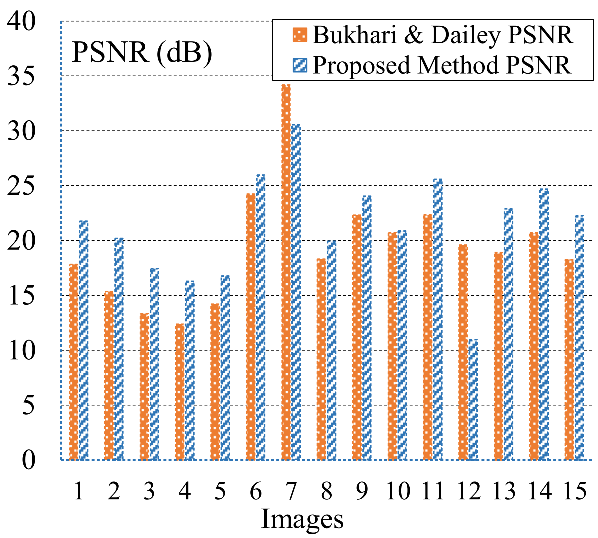


Fig. 3: RMSE and PSNR results for [16] and the proposed method. The rectified images are scaled and translated over a reasonable range and the best results are used.

The images rectified with either of the methods are compared with the ground truth images using RMSE and PSNR. The images are scaled and translated to produce the best result. Refer to Fig. 3 and Table I for quantitative results. For a general qualitative analysis of the methods, see Fig. 4. The borders of the images are especially useful in giving a sense of the estimated distortion characteristics. The proposed method estimates the distortion characteristics similarly for different images, even though the scenes are different. Average running time for the proposed method for a single 800×600 resolution image is 22.2 seconds. For scenes where lines are very dominant, the feature selection step described in Section II-D can be omitted. Then, the average running time for a single image would be 0.67 seconds.

TABLE I: Average RMSE and PSNR values for all images.

	Bukhari and Dailey [16]	Proposed Method
RMSE	45.06	31.53
PSNR (dB)	19.51	21.35

IV. Conclusion

The proposed method uses an error function that represents the amount of distortion well. Furthermore, due to iterative rectification of lines, the forwards distortion parameters are estimated, rather than estimating inverse parameters and approximating as a last step. Sequential backwards selection is applied to ensure robustness against line groups that emerge from non-linear structures. Contrary to using RANSAC, this results in a fully deterministic algorithm. The results show that the method works successfully in cluttered environments, if there are lines that represent the distortion characteristics.

References

[1] Roger Y Tsai, "A versatile camera calibration technique for high-accuracy 3d machine vision metrology using off-the-shelf tv cameras and lenses," *Robotics and Automation, IEEE Journal of*, vol. 3, no. 4, pp. 323–344, 1987.

[2] Juyang Weng, Paul Cohen, and Marc Herniou, "Camera calibration with distortion models and accuracy evaluation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 14, no. 10, pp. 965–980, 1992.

[3] Zhengyou Zhang, "A flexible new technique for camera calibration," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 22, no. 11, pp. 1330–1334, 2000.

[4] Robert T Collins and Yanghai Tsin, "Calibration of an outdoor active camera system," in *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, 1999.

[5] Olivier D Faugeras, Q-T Luong, and Stephen J Maybank, "Camera self-calibration: Theory and experiments," in *Computer Vision, European Conference on*, 1992, pp. 321–334.

[6] Andrew W Fitzgibbon, "Simultaneous linear estimation of multiple view geometry and lens distortion," in *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, 2001, vol. 1, pp. I–125.

[7] Srikumar Ramalingam, Peter Sturm, and Suresh K Lodha, "Generic self-calibration of central cameras," *Computer Vision and Image Understanding*, vol. 114, no. 2, pp. 210–219, 2010.

[8] Moumen Ahmed and Aly Farag, "Nonmetric calibration of camera lens distortion: differential methods and robust estimation," *Image Processing, IEEE Transactions on*, vol. 14, no. 8, pp. 1215–1230, 2005.

[9] Frederic Devernay and Olivier Faugeras, "Straight lines have to be straight," *Machine Vision and Applications*, vol. 13, no. 1, pp. 14–24, 2001.

[10] Christian Brauer-Burchardt and Klaus Voss, "A new algorithm to correct fish-eye-and strong wide-angle-lens-distortion from single images," in *Image Processing, International Conference on*, 2001, vol. 1, pp. 225–228.

[11] Thorsten Thormählen, Hellward Broszio, and Ingolf Wassermann, "Robust line-based calibration of lens distortion from a single view," in *Proceedings of MIRAGE*, 2003, pp. 105–112.

[12] John Mallon and Paul F Whelan, "Precise radial un-distortion of images," in *Pattern Recognition, International Conference on*, 2004, vol. 1, pp. 18–21.

[13] Duane C Brown, "Close-range camera calibration," *Photogramm. Eng*, vol. 37, pp. 855–866, 1971.

[14] Diego Gonzalez-Aguilera, Javier Gomez-Lahoz, and Pablo Rodríguez-Gonzálvez, "An automatic approach for radial lens distortion correction from a single image," *Sensors Journal, IEEE*, vol. 11, no. 4, pp. 956–965, 2011.

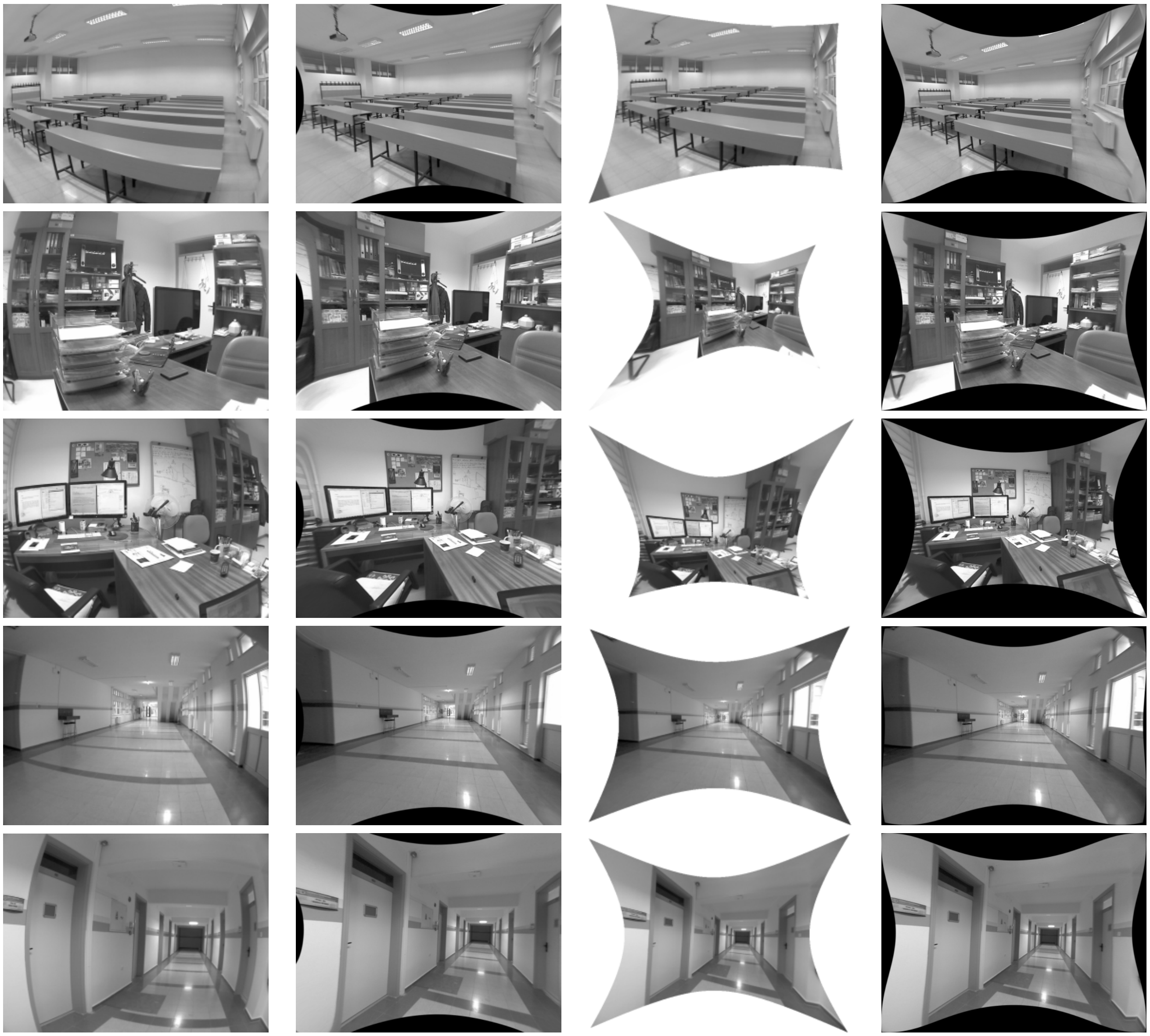


Fig. 4: First column is the original image, second column is rectified by pre-calibration parameters, third column is Bukhari and Dailey's results [16], fourth column is the proposed method's results. The indices of the images are 15, 4, 5, 7 and 11, in this particular order.

[15] Lazaros Grammatikopoulos, George Karras, and Elli Petsa, "An automatic approach for camera calibration from vanishing points," *ISPRS journal of photogrammetry and remote sensing*, vol. 62, no. 1, pp. 64–76, 2007.

[16] Faisal Bukhari and Matthew N Dailey, "Automatic radial distortion estimation from a single image," *Journal of mathematical imaging and vision*, vol. 45, no. 1, pp. 31–45, 2013.

[17] Luis Alvarez, Luis Gómez, and J Rafael Sendra, "An algebraic approach to lens distortion by line rectification," *Journal of Mathematical Imaging and Vision*, vol. 35, no. 1, pp. 36–50, 2009.

[18] Rita Cucchiara, Costantino Grana, Andrea Prati, and Roberto Vezzani, "A hough transform-based method for radial lens distortion correction," in *Image Analysis and Processing, International Conference on*, 2003, pp. 182–187.

[19] Rickard Strand and Eric Hayman, "Correcting radial distortion by circle fitting," in *BMVC*, 2005.

[20] Moumen T El-Melegy and Aly A Farag, "Can lens distortion be calibrated from an image of a smooth, textureless lambertian surface?," in *Image Processing, International Conference on*, 2012, pp. 321–324.

[21] C Brauer-Burchardt and Klaus Voss, "Automatic correction of weak radial lens distortion in single views of urban scenes using vanishing points," in *Image Processing, International Conference on*, 2002, vol. 3, pp. 865–868.

[22] Cuneyt Akinlar and Cihan Topal, "Edlines: Real-time line segment detection by edge drawing (ed)," in *Image Processing, International Conference on*, 2011, pp. 2837–2840.

[23] Burak Benligiray and Cihan Topal, "Lens distortion rectification using triangulation based interpolation," in *Advances in Visual Computing, International Symposium on*, 2015, pp. 35–44.

[24] Jeffrey C Lagarias, James A Reeds, Margaret H Wright, and Paul E Wright, "Convergence properties of the nelder–mead simplex method in low dimensions," *SIAM Journal on optimization*, vol. 9, no. 1, pp. 112–147, 1998.

[25] J. Y. Bouguet, "Camera calibration toolbox for Matlab," .