Camera Autocalibration using Predominantly Planar Aerial Imagery

Team:	Marta Palomar	Carlos Querejeta	Michael Hardt
SUNet ID:	mpalomar	carlosq	mwhardt

1. Introduction

Monocular camera systems are more and more being used for positioning in UAVs (*Unmanned Aerial Vehicles*) due to their low size, weight and power. However, pose estimation algorithms based on a single camera are extremely sensitive to any error in the camera's intrinsic parameters. Laboratory calibration techniques can provide estimations of the camera intrinsics that may not hold during high-altitude flight. A calibrated camera can also experience certain degradation during operations (caused by vibrations, temperature changes, varifocals [1]), thus leading to small shifts in the optical parameters. To ensure accuracy in position estimation, it is crucial to have a real-time update of the camera intrinsics.

2. Problem Statement

The project objective is to address the autocalibration problem using aerial imagery with a near planar terrain in view. A specific algorithm is implemented and described in the Technical Approach section. Different parallel, experimental validation approaches shall be taken by the members of the team using different datasets. One dataset consists of runway images taken from an airplane during a landing procedure. A close approximation of its camera calibration and accurate ground truth position and orientation data are known. A second dataset consists of transformed satellite imagery to create a virtual aerial view from a specified pinhole camera calibration. A third dataset consists of lab tests with a digital camera. In all cases, the camera intrinsics and distortion model are known beforehand, so this information can be used to evaluate the results. After these initial investigations, the team shall combine their efforts to apply the solution to real aerial imagery without known landmarks in view. These shall depend upon a further dataset with a downward facing camera and known calibration for verification purposes.

3. Technical Approach

Several publications [2–6] cover methods for self-calibration and known-target camera calibration. *Herrera et al*. [7] propose a novel approach for self-camera calibration using a planar scene with unknown texture which claims to achieve similar accuracy to checkerboard-based calibration approaches. This project is primarily based upon this latter work while exploring its application to aerial imagery. The considered camera model is a pinhole camera with nonsquare pixels, radial distortion and zero skew. The distortion function

Figure 1. Image taken during landing

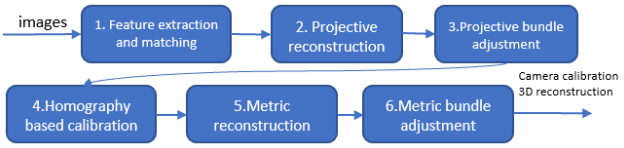

Figure 2. Steps of the algorithm by *Herrera et al*. [7]

$$p = (p_n - p_0)\left(1 + ||p_n - p_0||^2 d_0 + ||p_n - p_0||^4 d_1\right) + p_0 \quad (1)$$

has two distortion parameters d_0, d_1, where p is a point in the image, p_0 is the principal point, and p_n is the projected point. The resulting model thus has 6 *dof*: focal lengths f_x, f_y, principal point $p_0 = (u_0, v_0)$, and distortion coef's $d = (d_0, d_1)$.

The method consists in six steps which are depicted in figure 2: 1) *Feature extraction and matching*, 2) *Projective Reconstruction*, 3) *Projective bundle adjustment*, 4) *Homography based calibration*, 5) *Metric reconstruction*, and 6) *Metric bundle adjustment*.

3.1. Feature extraction and matching

The input to the algorithm is a sequence of images of a fundamentally planar scene. For each image, feature points are identified using SIFT [8] followed by a FLANN [9] matching algorithm applied to all images with respect to the first one. The first image is arbitrarily considered as the reference. Images with few correspondences and feature points that fall outside of the planar structure of the scene are discarded through an additional filtering stage.

3.2. Projective Reconstruction

The purpose of this step is to make a first estimation of the feature point world coordinates and the perspective transformation to their corresponding observations. As the points in the scene belong to a planar structure, these can be defined by 2D coordinates as $y = [x, y]^\top$. The world reference frame can be arbitrarily set in the center of the reference camera frame resulting in the initial world coordinates being the same as their observations in the reference image $y = p_0$. The objective of this step is then limited to finding the homographies H_i that transform observations in the reference image to their correspondences in all other images. In order to reduce the effect of outliers a RANSAC method is applied.

3.3. Projective Bundle Adjustment

Distortion effects haven't been accounted for yet and can result in significant error. Additionally, the homographies have been computed in a pairwise manner using the reference image which ignores the fact that the other images might constrain each other's projective pose. This step estimates the camera distortion (d_0, d_1) and central point (u_0, v_0) as well as refine the homography matrices H_i and the world coordinates y. A bundle adjustment process is conducted consisting in a nonlinear minimization of the following function: 2.

$$\underset{d',p_0,H_i,y_k}{\operatorname{argmin}} \sum_{i}^{n_{img}} \sum_{k}^{n_{pts}} \rho\left(||p_{ik} - D(H_i y_k)||^2\right) + \lambda||p - p_0||^2 \quad (2)$$

where d', p_0, H_i, y_k are respectively the distortion parameters, principal point, homography matrices, and point world coordinates. $D()$ is the distortion function of expression (1) and ρ is a robust function to reduce the influence of outliers. Experiments show slightly better results using a *Cauchy* function as opposed to a linear one. $\lambda||p - p_0||^2$ is a regularization term to bias the central point towards the center of the image, where λ is a hyper-parameter.

3.4. Homography based Self-Calibration

The purpose of this step is to recover the camera's focal length f_x and f_y as well as a better estimation of its central point $p_0 = (u_0, v_0)$. This step does not consider distortion since it has been already been corrected in the previous stage.

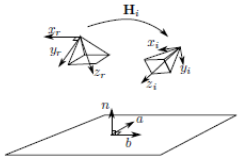

Figure 3. Planar self-calibration geometry

A set of constraints is derived based on the assumption of the planar scene. The scene plane in world coordinates can be defined by its 3D normal vector $\boldsymbol{n_0}$ and two orthogonal basis vectors which span the plane $\boldsymbol{a_0}$ and $\boldsymbol{b_0}$ as depicted in figure 3. As a result, the orthogonality and equal norm constraints can be expressed as follows:

$$\boldsymbol{a}_i^\top \boldsymbol{b}_i = 0; \qquad \boldsymbol{a}_i^\top \boldsymbol{a}_i - \boldsymbol{b}_i^\top \boldsymbol{b}_i = 0 \quad (3)$$

These vectors may be readily derived for the reference image and derived from the remaining images using: ($\boldsymbol{e}$ is an arbitrary auxiliary fixed vector that is not parallel to $\boldsymbol{n}_0$)

$$\boldsymbol{a}_0 = \boldsymbol{n}_0 \times \boldsymbol{e}; \qquad \boldsymbol{b}_0 = \boldsymbol{n}_0 \times \boldsymbol{a}_0 \quad (4)$$

$$\boldsymbol{a}_i = \boldsymbol{K}^{-1}\boldsymbol{H}_i\boldsymbol{K}\boldsymbol{a}_0; \qquad \boldsymbol{b}_i = \boldsymbol{K}^{-1}\boldsymbol{H}_i\boldsymbol{K}\boldsymbol{b}_0; \quad (5)$$

These are the constraints that enable the recovery of the camera's focal lengths and central point embedded in $\boldsymbol{K}$. In order to use a nonlinear minimization method to obtain both $\boldsymbol{K}$ and $\boldsymbol{n}_0$, an initial guess for the focal length is needed. This is accomplished using the following simplifications: i) known plane normal $\boldsymbol{n}_0 = [0, 0, 1]$ (scene plane fronto-parallel to the reference image), ii) $\boldsymbol{K} = diag(f, f, 1)$, and using equations 3 and 5 we obtain:

$$h_{31}h_{32}f^2 + h_{11}h_{12} + h_{21}h_{22} = 0 \quad (6)$$

$$f^2h_{31}^2 - f^2h_{32}^2 + h_{11}^2h_{21}^2 - h_{12}^2 - h_{22}^2 = 0 \quad (7)$$

Once an initial guess for the focal length is estimated, a nonlinear minimization procedure can be conducted to obtain $\boldsymbol{K}$ and $\boldsymbol{n}_0$ as follows:

$$\underset{\boldsymbol{K},\boldsymbol{n}_0}{\operatorname{argmin}} \sum_{i}^{n_{\text{images}}} \frac{\boldsymbol{a}_i\boldsymbol{b}_i}{||\boldsymbol{a}_i||||\boldsymbol{b}_i||} + \sum_{i}^{n_{\text{images}}} \left(1 - \frac{||\boldsymbol{b}_i||^2}{||\boldsymbol{a}_i||^2}\right) \quad (8)$$

The result of this step of the method is an estimation of the camera's intrinsic parameters in $\boldsymbol{K}$ and the normal to the scene plane $\boldsymbol{n}_0$

3.5. Metric Reconstruction

With the initial estimate of the cameras' intrinsics, it is possible to compute the extrinsics by assuming an approximation for the first (reference) frame camera extrinsics. A new world reference frame is taken such that the scene plane is placed at $z = 0$. The estimate for the rotation and translation of the reference camera is given by the relations:

$$\boldsymbol{R} = [\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{n}]^T \qquad \boldsymbol{t} = -\boldsymbol{R}\boldsymbol{n} \tag{9}$$

The next step is to triangulate all the feature points in the reference frame in order to obtain the 3D world coordinates of points $\mathbf{X}$. This is achieved by intersecting the light ray that goes from the center of the reference camera to each feature point with plane $z = 0$. First, we take the 2D coordinates $\boldsymbol{x}_d$ of all feature points (in homogeneous coordinates) and revert the projection (multiplying by $\boldsymbol{K}^{-1}$) in order to obtain their coordinates in the camera's reference system $\boldsymbol{x}_n$. Then the direction of each ray $\boldsymbol{l}$ can be computed as:

$$\boldsymbol{l}_x = \boldsymbol{R}^T \boldsymbol{x}_n \tag{10}$$

And the intersection with plane $z = 0$ follows the expression:

$$\boldsymbol{X} = \frac{\boldsymbol{l}(d - \boldsymbol{b} \cdot \boldsymbol{n})}{\boldsymbol{l} \cdot \boldsymbol{n}} + \boldsymbol{b} \tag{11}$$

In this case, $d = 0$, $\boldsymbol{n} = [0, 0, 1]^T$ and $\boldsymbol{b}$ stands for the center of the reference camera. By solving equation 11 with the data from the reference frame, we obtain the 3D positions of the reference feature points $\boldsymbol{X}_r$.

This gives us the full metric reconstruction for the reference frame. For the rest of keyframes, provided that K is known, the extrinsics of the cameras are obtained by solving the Perspective n-Point problem, that requires as inputs the 2D pixel coordinates of corresponding points in each keyframe and their analogous in $\boldsymbol{X}_r$. By completing this procedure, the full metrics of each camera and frame are known: the 3D points in the scene, the intrinsic and extrinsic camera parameters.

3.6. Metric Bundle Adjustment

The purpose of this final step is to simultaneously refine the 3D coordinates of the points in the scene, the intrinsic camera parameters including distortion and the camera poses for each image. To that end a non-linear optimization procedure is applied to minimize the reprojection errors of 3D points as follows:

$$\underset{\boldsymbol{K}, \boldsymbol{d}, \boldsymbol{R}_i, \boldsymbol{t}_i, \boldsymbol{x}_k}{\operatorname{argmin}} \sum_{i}^{n_{images}} \sum_{k}^{n_{points}} \rho\left(||\boldsymbol{p}_{ik} - P_i(\boldsymbol{x}_k)||\right)^2 \tag{12}$$

4. Final Results

This section presents the results obtained from the application of the method to images obtained from: (i) close range (ii) aerial imagery and iii) satellite imagery. An implementation of the method has been developed in Python from scratch and is available at [10], making use of `opencv` (feature detection and matching, initial homography estimation) and `scikit` (nonlinear and least squares solvers) libraries. The implementation includes all six steps of the method detailed in section 3

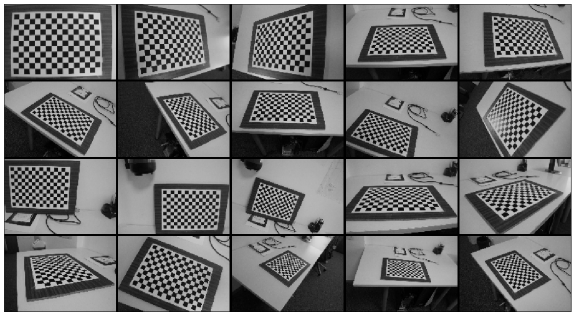

Figure 4. Checkerboard images used for baseline calibration

Figure 5. Poster images used to apply the planar self-calibration method

4.1. Close range imagery

One of the outcomes of the milestone report was that the camera used for close range imagery applied unwanted distortion correction to the images. We have repeated this experiment with a Canon® EOS-450D DSLR camera with 18-50mm lenses ensuring that no distortion correction is applied to the images. We chose the wide end of the lenses (18 mm) as it shows the most distortion, resulting in a more challenging problem for the self-calibration method. A major difference with respect to the experiment conducted in the milestone, is that the missing *metric reconstruction (3.5)* and *final bundle adjustment (3.6)* steps are now implemented, which results in a dramatic improvement of the camera calibration results.

A baseline calibration for the camera has been initially obtained using the well known *Camera Calibration Toolbox for Matlab®* [11]. To this end a set of 20 pictures of a known checkerboard pattern[1] were taken (see figure 4), and then resized to 800×533 pixels. In order to apply the planar self-calibration method, a sequence of 37 images of a poster (planar surface with unknown texture) were taken at close range in different poses, figure 5 shows a subset of such images.

Table 1 compiles the results obtained for the camera's

[1]The checkerboard pattern is made up by a grid of 14x9 squares of 25 mm length

focal length, principal point, and distortion parameters as well as the reprojection error of the $3D$ scene points onto the images as follows:

- the *baseline* column shows the results obtained with the Matlab calibration toolbox applied to the set of checkerboard images
- the *planar self-calibration* column shows the results obtained by applying the method detailed in this project to the sequence of poster images.
- the *error* column shows the deviation betweeen the results obtained by the planar self-calibration and the baseline checkerboard-based calibration methods. For evaluation purposes, inside the parentheses the result reported in the project milestone is shown, when the method implementation was not completed yet.
- the *published result* column shows the results published in the reference paper [7]

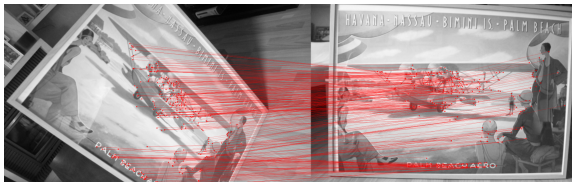

Figure 6. Feature points and matches concentrated towards the center of the image

It can be observed that the focal length obtained by the planar self-calibration method is very accurate with a deviation of only 1.12% with respect to the baseline and really close to the nominal $18mm$ focal length, even though it is slightly greater than the one published in [7] we believe this minor difference could be attributed to the different camera and planar surface images being used in the publication. It has to be highlighted the dramatic improvement obtained with respect to the milestone results, which emphasizes the effect of the *metric reconstruction* and *final bundle adjustment* steps. It is also relevant to note that these focal length results capture very well the squared pixel structure (deviation in x and y very similar) whereas in the previous report the focal length deviation between both dimensions was significant (9.1% vs 6.9 %).

Analyzing the principal point results we observe that they are on par with those obtained in the milestone but still not as good as those published in the reference paper. Again, the camera and images used can play a significant role here, we will show better results for the principal point when the very same method uses aerial imagery and a different camera in section 4.2.

	baseline	**planar self-calibration**	**error %**	**published result %**
f_x	657.61 px (18.25 mm)	664.95 px (18.45 mm)	1.12 (9.1)	0.69
f_y	658.27 px (18.27 mm)	666.04 px (18.49 mm)	1.18 (6.9)	0.93
u_0	401.66	425.68	5.9 (2.1)	2
v_0	289.90	283.17	2.3 (4.7)	-0.37
d_0	-0.176	$-3.75 \cdot 10^{-7}$	100 (100)	-37.63
d_1	0.104	$5.17 \cdot 10^{-13}$	100 (100)	-14.45
error	0.65	0.58	-	1.94 px

Table 1. Close range imagery self-calibration method results

The reprojection error obtained by the planar calibration method is of only $0.58\ px$, much better than those published in the reference paper ($1.94\ px$) and slightly outperform those of the baseline calibration ($0.65\ px$), which is unusual. Such good results could be attributed to a higher density of feature points concentrated around the central part of the image in most of the samples, where the effect of the distortion is not substantial, this effect can be observed in figure 6. This seems to be corroborated by the results obtained in the distortion parameters which are very small in magnitude, unveiling that just a slight distortion is needed in order to obtain a good reprojection error. It has to be noted that the reprojection error drives the *metric bundle adjustment* step from which the final calibration results are obtained. These good results could also be attributed to the published method using a faster ORB feature detector as opposed to SIFT selected in this project which identifies more features in a more homogeneous distribution throughout the image resulting in better homography estimations.

One major outcome of the analysis of close range imagery is that the planar scene used for self-calibration has to be carefully selected so that its distribution of features is homogeneous throughout the scene, otherwise as has been demonstrated, the results of the distortion model are poor. Anyway, the method works very well for the remaining intrinsics of the camera, especially for the focal length.

4.2. Aerial Imagery

The implemented algorithm has been tested using real aerial imagery from a landing maneuver at Sanderson Field Airport (Figure 7). The images were taken with a forward-looking GoPro camera at a resolution of 4K (3840x2160 px). The full set intrinsic parameters of the camera were also given by the image provider and will be employed for validation. It is noticeable that the input images do not have a significant amount of texture in order to enhance feature extraction. However, in the scene region closer to the runway, our algorithm is able to obtain good matches between

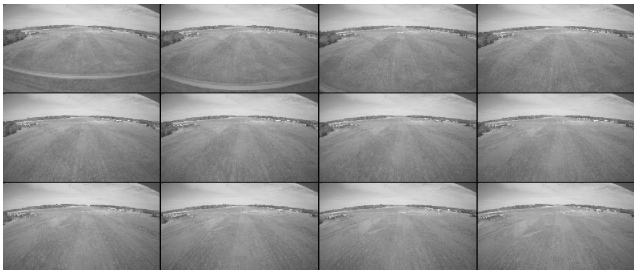

Figure 7. Input aerial images to the implemented autocalibration tool

	Ground truth	Self-calibration	error (%)
f_x	1749.42 px	1766.25 px	0.94
f_y	1747.93 px	1747.92 px	0.98
u_0	1891.50 px	1891.50 px	1.3
v_0	1085.92 px	1085.92 px	-0.07
d_0	-0.268	-0.206	23.1
d_1	0.0001	0.011	-

Table 2. Real camera intrinsics compared to the estimation with the self-calibration algorithm.

all the images in the sequence, as it can be appreciated in Figures 8 and 9.

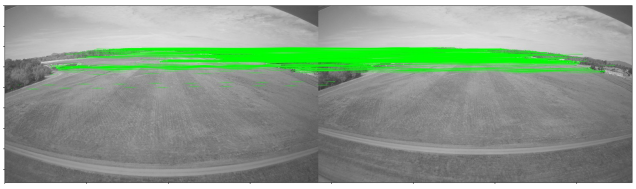

Figure 8. Matching between real aerial images with the self-calibration algorithm

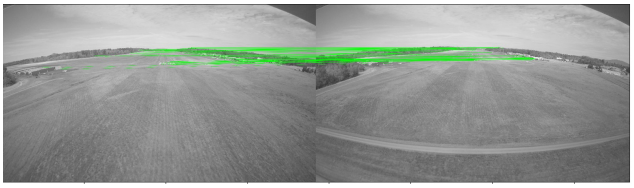

Figure 9. Found correspondences between the last frame of the sequence and the reference frame

A comparison between the intrinsic parameters estimated with the proposed self-calibration methodology and the real (ground truth) intrinsics is given in Table 2. Also, the two radial distortion parameters included in the implemented tool have been computed.

It is noteworthy that all intrinsic parameters are accurately estimated, with error values lower than 1% with respect to the ground truth. These results are encouraging as it was expected that the strong perspective that is characteristic of landing imagery and the lack of texture in most of the scene would lead to a worse approximation. If compared to the results obtained at the halfway of the project (milestone report), it is clearly noticeable that the implementation of the metric bundle adjustment has led to a great reduction of estimation errors in all intrinsic parameters, but specially in the principal point.

On the other hand, a good attempt has been made to capture the distortion of the images. The lenses employed during the landing did present both radial and tangential distortion, and these were defined by means of a 5-parameter distortion model. Given that the algorithm implemented in this project does only account for radial distortion, it was challenging to obtain a fair approach given the disparity between models. Nonetheless, the estimate of the first radial distortion term d_0 is the same order of magnitude than the real one for the lenses, with an error level of 23%. This difference is expected to come from the fact that the extracted features for matching are not homogeneously distributed along the entire scene.

4.3. Satellite Imagery

As an approximation to the objective of autocalibrating at altitude, the approach is taken here to transform satellite imagery to a this has the advantage that a perfect ground truth and intrinsics are known, and the method can be easily tested.

A set of 10 images are used representing a virtual, downward-facing camera from a plane flying at approximately 900*m* above the terrain. The plane is performing a left-handed turn during the image sequence, and in addition, small angular variations are reflected which may occur from control and turbulence. Thus, small motions exist in all 6-*dof*, yet the primary motion is parallel to the underlying terrain. The images are generated using the application `osgEarth` which transforms satellite imagery obtained from [12] to a desired viewpoint and perspective transformation. The generated imaes are shown in Figures 10 and 11.

The described algorithm produced erratic results depending upon the initial conditions provided. An in-depth analysis of the data indicated an accurate feature matching with subpixel reprojection errors. However, some of the initially estimated homography matrices have condition numbers on the order of 10^6 which is poor.

In order to improve the optimization, a number of measures have been taken and implemented:

- selected feature points are scaled such that their mean square distance to the origin is 2 pixels

- optimize over the homography inverse in the bundle

Figure 10. Aerial view generated with osgEarth from satellite image.

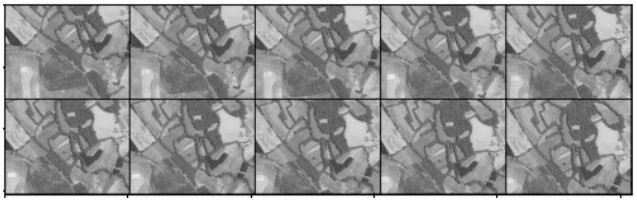

Figure 11. Sample aerial images used for autocalibration.

	Ground truth	**Self-calibration**	**error (%)**
f_x	1117.65 px	1527.25 px	36.7
f_y	1117.65 px	1455.30 px	30.3
u_0	511.50 px	480.65 px	6.0
v_0	383.50 px	415.15 px	8.3

Table 3. Real camera intrinsics compared to the estimation with the self-calibration algorithm.

adjustment such that the inverse operation doesn't enter into the optimization

- the homographies and corresponding image pairs with the best condition numbers are selected
- explicit Jacobians are programmed for the projective and metric bundle adjustments within the autocalibration pipeline

Finally, the best results obtained are shown in Table 3 which are not promising. After further investigation and literature review, it has been identified that the case of pure translation is a further degenerate case within the already degenerate case we are studying. Considering the high altitude and the relatively small rotation angles between sample images, it is hypothesized that this experiment is too close to this particular case. Zhang demonstrated the degeneracy of a pure translation in which the planar target in one image is parallel to a second one in [13]. In this case, the addition of that image's homography does not add any more constraints to the optimization as they are linear dependent upon the existing ones. Zhang proposed an alternative algorithm specifically for this case which, however, required the knowledge of the precise translation vector between frames which is quite restrictive.

Yang [14] recently published an improvement to this autocalibration procedure which only requires knowledge of the distance between parallel frames. A subsequent parallel image with this approach adds one further constraint upon the absolute conic $\omega = (K\,K^T)^{-1}$ based upon the inter-homography matrix and the relative distance between images. Further necessary assumptions are that the instrinsics are skewless and the ratio of focal length is known, i.e. $\frac{f_x}{f_y}$. Up to four constraints can be obtained by this means upon the absolute conic, which when combined with the two original constraints upon the reference image and the planar scene gives the six constraints necessary to solve for the absolute conic. A Cholesky factorization can then be used to solve for the intrinsic matrix K. This value could then be used as an input in the metric reconstruction bundle adjustment step described in Section 3.6.

5. Conclusions

One of the main conclusions that can be identified is that the type of planar scene strongly influences the results of the presented self-calibration method. Even in a controlled environment as detailed in section 4.1, the planar scene has to be carefully selected so that feature points are homogeneously distributed throughout the scene, otherwise the method will fail to identify a representative distortion model. One potential solution could be to further enhance the feature extraction process to promote such homogeneous distribution of features, or at least notify when it fails to do so. Other than that, the results for the remaining camera intrinsics in a controlled environment are really promising, and rival those obtained by means of conventional checkerboard patterns and manual corner identification, but with a more convenient setup and no manual intervention.

The presented self-calibration method has been tested using real aerial images from a landing maneuver. The estimated intrinsic parameters reveal a reduced error level with respect to the real ones, which suggests the suitability of this proposed methodology for camera self-calibration prior to autolanding vision-based maneuvers. It is expected that using images with more texture and more evenly distributed features the radial distortion parameters of real lenses could be reasonably estimated for cases in which the tangential components of distortion have little influence over the radial ones.

Finally, for aerial views at a significant altitude above

the terrain, the investigated method delivered poor results not being able to converge to the true, known calibration. It is estimated that the cause is that the relative pose between sample images is too close to a pure translation which is a degenerate case for this approach. A suitable approach has been identified which serves to calculate a good, initial value for the intrinsics for this particular case, and which may subsequently be used in the final metric reconstruction step for further refinement.

References

[1] K. Celik and A. Somani, "Wandless realtime autocalibration of tactical monocular cameras," in *Proceedings of the International Conference on Image Processing, Computer Vision, and Pattern Recognition (IPCV)*, p. 1, The Steering Committee of The World Congress in Computer Science, Computer . . . , 2012.

[2] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*. USA: Cambridge University Press, 2 ed., 2003.

[3] S. Bougnoux, "From projective to euclidean space under any practical situation, a criticism of self-calibration," *Sixth International Conference on Computer Vision (IEEE Cat. No.98CH36271)*, pp. 790–796, 1998.

[4] B. Triggs, "Autocalibration from planar scenes," in *European Conference on Computer Vision (ECCV '98)* (H. Burkhardt and B. Neumann, eds.), vol. 1406 of *Lecture Notes in Computer Science (LNCS)*, (Freiburg, Germany), pp. 89–105, Springer-Verlag, June 1998. Work supported by Esprit LTR project Cumuli.

[5] B. Bocquillon, P. Gurdjos, and A. Crouzil, "Towards a guaranteed solution to plane-based self-calibration," in *ACCV*, 2006.

[6] P. Gurdjos and P. Sturm, "Methods and geometry for plane-based self-calibration," in *2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings.*, vol. 1, pp. I–I, 2003.

[7] D. H. C., J. Kannala, and J. Heikkilä, "Forget the checkerboard: practical self-calibration using a planar scene," *IEEE Winter Conference on Applications of Computer Vision (WACV)*, 2016.

[8] D. G. Lowe, "Object recognition from local scale-invariant features," in *Proceedings of the Seventh IEEE International Conference on Computer Vision*, vol. 2, pp. 1150–1157 vol.2, 1999.

[9] M. Muja and D. G. Lowe, "Fast approximate nearest neighbors with automatic algorithm configuration," in *In VISAPP International Conference on Computer Vision Theory and Applications*, pp. 331–340, 2009.

[10] Palomar-Querejeta-Hardt, "Winter 2021 cs231a project implementation." `https://github.com/txoritxo/cs231a_Project`.

[11] J.-Y. Bouguet, "Camera calibration toolbox for matlab." `http://www.vision.caltech.edu/bouguetj/calib_doc/`.

[12] `https://arcgisonline.com`.

[13] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 22, pp. 1330–34, 2000.

[14] M. Yang, X. Chen, and C. Yu, "Camera calibration using a planar target with pure translation," *Applied Optics*, vol. 58, pp. 8362–70, 2019.