CALIBRATION AND DISTORTION COMPENSATION ALGORITHMS OF DIGITAL CAMERAS USED IN UAV

Teodora Petrova, Marin Marinov and Zhivo Petrov

Trakia University, Stara Zagora
Georgi Benkovski Bulgarian Air Force Academy, Dolna Mitropolia
Rakovski National Defence College, Sofia

The paper presents algorithms for calibrating and correcting the distortions of a digital photo camera on board the UAV in order to reduce the distortions in the obtained images, which provides a more accurate analysis of details. A model of the main internal parameters of the camera is presented, supplemented with the distortion coefficients, and a procedure for external calibration and self-calibration of the camera to determine the optimal values of these parameters. A genetic algorithm is proposed for minimizing the objective function, image correction, and calibrating and correcting camera distortions. The results show that using the proposed algorithm, the reduction of pixel bias is more than 100 times, and the obtained estimates of the internal parameters are close to the real ones and can be used to correct the distortions

In recent decades, images from cameras on board unmanned aerial vehicles (UAVs) have increasingly been used to monitor the Earth's surface, including vegetation. Digital processing methods have found wide applications in industry, art, medicine, space, ecology, etc. They are applied in managing automated processes for detecting and tracking objects, image recognition, and environmental monitoring. Image formation, enhancement of their quality, and automated processing of aerial images, including images taken by satellites, unmanned aerial vehicles, and synthetic aperture radars, are subjects of numerous researches and developments. Automated analysis is widely applied in remote monitoring systems for localities, forests, calculating the area of damage, crop monitoring, reconnaissance, and fire safety systems.

In aerial photography, each object, camera, and aircraft has its coordinate system. All coordinate systems are right-oriented. These coordinate systems facilitate the description of projective cameras.

When developing camera calibration models, it is necessary to take into account the successive transition from one coordinate system to another.

The aim of this study is to synthesize algorithms for calibration and correction of distortions of a camera on board the UAV in order to reduce the distortions in the obtained images, which will allow a more accurate analysis of the details on them.

Mathematical model for the calibration of an aerial camera.

The physical parameters of the camera are external and internal. Conversion object coordinates to center camera frame coordinates is necessary to use external parameters of the camera. In UAV using more than one camera to characterize the connections between the cameras external parameters are used. Standard camera models are collinear. In these models, each point of the object is projected by a straight line into the projection centre trough the image plane. The projected centre (X_0, Y_0, Z_0) is the begining of the camera coordinate system and the *z*-axis of the camera frame is perpendicular to the image plane. Euler planar rotation angles *ω, φ* and, *k* which is explained as a series of three simple rotations about the *x, y* and *z*-axes are used as a rotation between the object and camera coordinate systems. The clockwise rotations is used, first around the *x*-axis, then around the *y*-axis, which has already been rotated once, and finally around the *z*-axis, which has already been rotated twice.

For an arbitrary point of the object *P*, with coordinates (X_i, Y_i, Z_i) in the image coordinate system, it is first necessary to switch to the camera coordinate system (x_i, y_i, z_i). The transformation consists of translation and rotation:

$$\begin{bmatrix} x_i \\ y_i \\ z_i \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{bmatrix} \begin{bmatrix} X_i \\ Y_i \\ Z_i \end{bmatrix} + \begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix}, \tag{1}$$

where:

$m_{12} = sin\,\omega\, sin\,\phi\, cos\,k - cos\,\omega\, sin\,k$ $\qquad m_{11} = cos\,\phi\, cos\,k$

$m_{22} = sin\,\omega\, sin\,\phi\, sin\,k + cos\,\omega\, cos\,k$ $\qquad m_{21} = cos\,\phi\, cos\,k$

$m_{13} = cos\,\omega\, sin\,\phi\, cos\,k + sin\,\omega\, sin\,k$ $\qquad m_{31} = -\,sin\,\phi$

$m_{23} = cos\,\omega\, sin\,\phi\, sin\,k - sin\,\omega\, cos\,k$ $\qquad m_{31} = sin\,\omega\, cos\,\phi$

$m_{33} = cos\,\omega\, cos\,\phi$

Internal camera parameters often contain effective focal length f, zoom factor S_u, and coordinates of the center of the image (u_0, v_0), also called the principal point. The begining of the image coordinate system is most often in the upper left corner Eq. (1). The unit of measurement of the image coordinate system is a pixel, and therefore it is necessary to use the coefficients D_u and D_v for the transition from metric units to pixels. These coefficients are obtained from the camera data and the frame mounting. Their exact values are not critical in the calibration process, as they depend linearly on the focal length f and scale factor S_u. When using a standard camera model, the point (x_i, y_i, z_i) from the camera coordinate system is transformed to the image plane by the expression:

$$\begin{bmatrix} \tilde{u}_i \\ \tilde{v}_i \end{bmatrix} = \frac{f}{z_i} \begin{bmatrix} x_i \\ y_i \end{bmatrix} \qquad (2)$$

The corresponding image coordinates (u_i, v_i) in pixels are obtained from the image metric coordinates $(\tilde{u}_i, \tilde{v}_i)$ by the following transformation:

$$\begin{bmatrix} u_i' \\ v_i' \end{bmatrix} = \begin{bmatrix} D_u s_u \tilde{u}_i \\ D_v \tilde{v}_i \end{bmatrix} + \begin{bmatrix} u_0 \\ v_0 \end{bmatrix} \qquad (3)$$

The standard camera model is only an approximation of the true camera projection. It is a useful model that allows to find the relationship between the coordinates of the object and the image, through simple mathematical calculations. It does not achieve high accuracy of the obtained results, for this it is necessary to use a more complex camera model. The standard camera model is extended with several components for the systematic distortion of image coordinates. The most common fix is for radial/axial lens distortion, which causes the actual image points to shift radially in its plane. Radial distortion is expressed by [3]:

$$\begin{bmatrix} \delta u_i^{(r)} \\ \delta v_i^{(r)} \end{bmatrix} = \begin{bmatrix} \tilde{u}_i\left(k_1 r_i^2 + k_2 r_i^4 \ldots\right) \\ \tilde{v}_i\left(k_1 r_i^2 + k_2 r_i^4 \ldots\right) \end{bmatrix} \qquad (4)$$

where k_1 and k_2 are the coefficients of the radial distortion and $r_i = \sqrt{\tilde{u}_i^2 + \tilde{v}_i^2}$.

The centres of curvature of the lens surfaces are not strictly collinear, resulting in another type of distortion, decentred distortion, which includes a radial and tangential component. The tangential distortion has the form:

$$\begin{bmatrix} \delta u_i^{(t)} \\ \delta v_i^{(t)} \end{bmatrix} = \begin{bmatrix} 2p_1 \tilde{u}_i \tilde{v}_i + p_2\left(r_i^2 + 2\tilde{u}_i^2\right) \\ p_1\left(r_i^2 + 2\tilde{v}_i^2\right) + 2p_2 \tilde{u}_i \tilde{v}_i \end{bmatrix} \quad (5)$$

where p_1 and p_2 are the coefficients of the tangential distortion.

Linear distortion correction is used for some distortions. This correction is only used if the image axes are not perpendicular. In most cases, the deviation is small and insignificant. Another deviation is a subtle distortion of the prism. It results from poorly designed and manufactured lenses, as well as imperfect assembly of individual camera elements. This type of distortion can be adequately modelled by coupling the thin prism to the optical system, resulting in an additional number of axial and tangential distortions.

A suitable camera model to perform calibration is by combining expressions (3) – (5) to include both axial and tangential distortions [3]:

$$\begin{bmatrix} u_i \\ v_i \end{bmatrix} = \begin{bmatrix} D_u s_u \left(\tilde{u}_i + \delta u_i^{(r)} + \delta u_i^{(t)}\right) \\ D_v \left(\tilde{v}_i + \delta v_i^{(r)} + \delta v_i^{(t)}\right) \end{bmatrix} + \begin{bmatrix} u_0 \\ v_0 \end{bmatrix} \quad (6)$$

In this model, the four main internal camera parameters (f, s_u, u_0, v_0) are supplemented with the distortion coefficients $k_1, \ldots, k_n, p_1, p_2$.

The external camera calibration procedure is used to determine the optimal values of these parameters based on an image of a known object in three-dimensional space. In the case of self-calibration, coordinates of points from the object are included in a group of unknown parameters.

CAMERA CALIBRATION ALGORITHM.

Camera calibration algorithms use a model described by Eqs. (2) – (6) and captured calibration objects to estimate the internal (physical) parameters of a particular camera. These algorithms are of two main groups – linear and non-linear.

Linear algorithms do not use iterations, but they have two main disadvantages. Lens distortion cannot be included, and therefore the distortion effect cannot be calculated accurately. Due to the non-iterative algorithm, the actual constraints on the intermediate parameters are not considered. Therefore, in the case of noise, the intermediate solution does not satisfy the constraints, and the accuracy of the final result is lo.

Various errors accumulate in the process of image formation. If the systematic parts of the errors are compensated, the random errors are assumed to be white Gaussian noise. In order to achieve the best estimates of the camera parameters, the difference between the described model and N number of observed images (U_1, V_i), where $i = 1, \ldots, N$ is minimized. In the case of Gaussian noise, the objective function to minimize is a sum of squared errors:

$$F = \sum_{i=1}^{N} \sum_{j=1}^{M} \left(U_{ij} - u_{ij}\right)^2 + \sum_{i=1}^{N} \sum_{j=1}^{M} \left(V_{ij} - v_{ij}\right)^2 = = \sum_{i=1}^{N} \left[U_i - D_u s_u \left(\tilde{u}_i + \delta u_i^{(r)} + \delta u_i^{(t)}\right) + u_0\right]^2 + + \sum_{i=1}^{N} \left[V_i - D_v \left(\tilde{v}_i + \delta v_i^{(r)} + \delta v_i^{(t)}\right) + v_0\right]^2 \qquad (7)$$

Research in [3] shows that both types of distortions (axial or tangential) require two coefficients each. Linear distortion is negligible in modern CCD arrays and can be ignored.

Due to the nonlinear camera model, parameter estimation involves the application of an iterative algorithm. These algorithms are often slow to converge or may lead to finding a local minimum and calibration error. This drawback can be avoided by using parameters obtained from linear calibration methods as initial values for optimization. The general minimum is obtained after several successive iterations.

In this study, we propose using a genetic algorithm to minimize the objective function (as shown in Fig. 3). The first step of the algorithm is to determine and set the range within which to generate the population members of each generation. The next step is to generate the current population of 200 members (sets of values of the estimated parameters) with a uniform probability distribution for the first generation. The objective function is then calculated for each member of the population. The two sets of camera parameters for which the objective function has the smallest values are selected from the entire population. These members of the population are considered elite and will necessarily pass to the next generation without any changes.

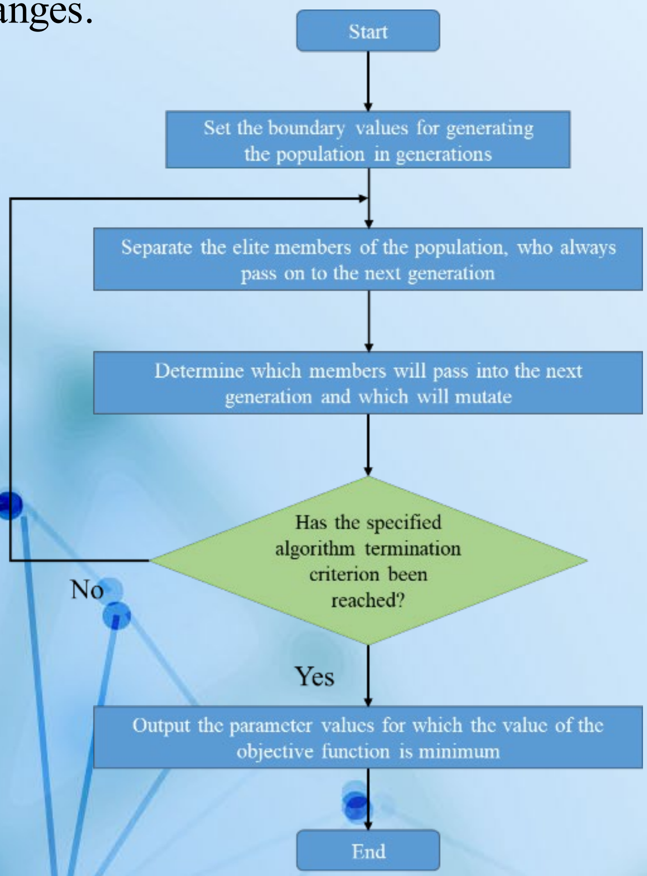


Fig. 3. Genetic algorithm.

In the next step, 20% of the population members are randomly selected to move to another population. The remaining 80% of the new population is obtained by crossing pairs of members of the current population. It is checked whether the criteria for ending the algorithm have been reached. If they are not, all the resulting members of the new generation obtained in the described ways are recorded as part of the new population. The main criterion for completing the algorithm is the condition that the cumulative change in the calculated objective function is less than 10^{-10}. In order to avoid the possibility of obtaining an infinite cycle, as an additional condition, an exit upon reaching 150 generations is set. After the algorithm is completed, the set of camera parameters of the last generation, for which the objective function has a minimum value, is the estimated parameter vector.

The number of external parameters depends on the number of observed images. With a known location of a three-dimensional reference object, an image of only one view of it is required. In the case of a planar (flat) reference object, the number of parameters that can be obtained from one of its views is limited. For this reason, more than one view is required to calculate all internal parameters. In this study, a black and white checkerboard surface with 6 rows and 8 columns, which is rotated about each axis about its center, is used as a reference flat object.

1. The following procedure is proposed to find the estimated parameters:
2. The image of the checkerboard in 7 different views is captured.
3. The Harris-Stephens algorithm is applied to detect corner pixels (61 per number of U_i and V_i coordinates) for each image.
4. From the known rotation angles, the true coordinates of the corner points of the checkerboard in the camera coordinate system for each view are found.
5. The genetic algorithm is applied, which searches for that set of camera parameters for which the function (7) is minimal.
6. The vector of estimated camera parameters $[f, S_u, u_0, v_0, k_1, k_2, p_1, p_2]^T$ obtained after the end of the genetic algorithm is stored.

Image correction algorithm.

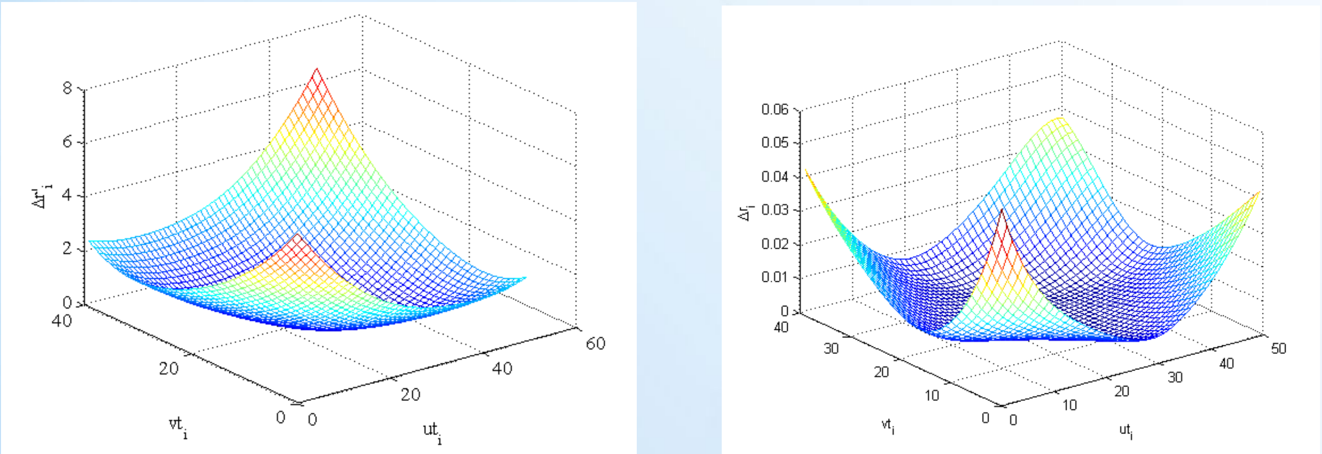
The camera model described by (6) gives the projection of the points of the three-dimensional space onto the image plane. But it does not provide a direct solution to the inverse projection problem by which to recover the line of sight from the image coordinates. If both axial and tangential distortion components are considered, then there is no analytical solution to the inverse transformation. The two axial distortion coefficients transform the camera model in Eq. (6) into a 5th-order polynomial.

Based on this mathematical model, an algorithm was developed to correct distortions in images:

1. A set of 2000 points$(\tilde{u}_i, \tilde{v}_i)$ is generated, evenly spaced over the entire image in 40 rows and 50 columns.
2. The internal camera parameters found from the calibration step are read.
3. Using equation (6), the distorted coordinates $(\tilde{u}'_i, \tilde{v}'_i)$ corresponding to the camera model are calculated.
4. The matrix T and the vector e are obtained.
5. The estimated parameters of the back-projection are found .
6. The distortions in the image are corrected.

To evaluate the algorithm for correcting distortions in images, a study was made for an image with dimensions of 1000×1200 pixels and camera parameters - $f = 15\,mm$, $S_u = 1.02$, $u_0 = 600$, $v_0 = 500$, $k_1 = 3$, $k_2 = 0.5$, $p_1 = 0{,}3$, $p_2 = 0.2$. The radial deviation of the pixels relative to the location of the pixels in a distortion-free image (ut_i, vt_i) was investigated:

Figure 4 shows the results for the uncorrected (Fig. 4 a) and corrected image (Fig. 4 b) for the 2000-point grid.



a) uncorrected image. b) corrected image.

Fig. 4. Pixel deviation for the uncorrected and the corrected image.

Figure 4 shows that for the mesh endpoints in the uncorrected image, the radial deviation reaches up to 8 pixels, while for the corrected image the deviation is below 0.06 pixels. The results show that using the proposed correction algorithm, the reduction in pixel bias is more than 100 times.

The study conducted on the synthesized algorithms for calibrating and correcting camera image distortions shows that they yield good results. By using these algorithms, the distortions in the captured images are significantly reduced, thus enabling a more precise analysis of the image details. There is no need to implement the image correction algorithm on board the UAV; thus, reducing the need for on-board equipment significantly.


THANK YOU
FOR YOUR ATTENTION