COBEM 2023 27TH INTERNATIONAL CONGRESS OF MECHANICAL ENGENEERING

ABCM Brazilian Society of Mechanical Sciences and Engineering
27th ABCM International Congress of Mechanical Engineering
December 04-08, 2023, Florianópolis-SC, Brazil

COB-2023-1894
A SIMPLIFIED STEREO COMPUTER VISION MODEL WITH FISHEYE CAMERAS TO DETERMINE THE HEIGHT OF CLOUDS OR DISTANT OBJECTS IN THE SKY

Darío Gerardo Fantini
Matheus Carvalho
Mario Benjamim Baptista de Siqueira
Energy and Environment Laboratory, Department of Mechanical Engineering, University of Brasilia, Brasilia, Federal District, Brazil
fantinidario@gmail.com, matheusjcarvalho@gmail.com, mariosiqueira@unb.br

Abstract. *Solar forecasting can be obtained through sky images by detecting clouds, responsible for the high-frequency variability of solar irradiation. Therefore, ceilometers are commonly used to perform this task, but are expensive. Thus, in this work we propose a model that uses stereo vision with fisheye cameras, with the objective of measuring the height of the cloud base. This system is currently installed in the Energy and Environment Laboratory of the University of Brasilia. The model proposed to determine the heights does not require the undistortion image process, typical to fisheye lenses, nor the rectification of the images, being these the major contributions compared to other existing models. For the validation of our model, a new validation method with a drone was used. With drone in stationary flight at various points and desired altitudes, we captured several images of the sky with our fisheye cameras. With this, we verified the heights obtained by our model and compared them with those reported by the drone. The first results showed an error of less than 10% (concerning the value reported by the drone) at heights up to 490 meters relative to the cameras.*

Keywords: *sky images, fisheye cameras, stereo vision model, solar forecasting.*

1. INTRODUCTION

In the literature different approaches to stereo computer vision are found, however, there are still few involving omnidirectional cameras as is the specific case of fisheye cameras. Abraham and Förstner (2005) presents an epipolar rectification model considering the equi-distance and stereographic model as fisheye projection models and use rectification of the images after undistortions (image flattening). Khomutenko et al. (2017) propose a method that is used to model fisheye cameras by projecting straight lines on conical sections, which will be the epipolar curves responsible for adapting the Semi-Global matching algorithm to stereo fisheye systems and finally, calculate a dense direct stereoscopic matching without fisheye distortion or rectification. Cai and Qiao (2021) bring a fisheye camera calibration method for stereo systems where they use the Levenberg-Marquarat algorithm to minimize the difference between 5 corresponding points of the two stereo images, considering both intrinsic and extrinsic parameters of both cameras. The aforementioned stereo calibration methods and also the one proposed in the OpenCV API require the use of a known pattern, such as the chessboard or dot pattern, where the dimensions of these are known. This is a similar procedure to the camera calibration presented above, but with the difference that, in this case, the chessboard or pattern of dots must be seen simultaneously by the cameras. In the recent work of Cao et al. (2022) proposes a stereo fisheye model of direct mapping of points in space, which is obtained using neural networks. This condition is difficult to meet here, since in the proposed system, the cameras will be at a distance greater than 40 meters. It should be emphasized that the proposed model does not use rectification or image flattening, which differentiates it from the mentioned methods.

The use of stereo vision with fisheye cameras applied to cloud base height measurement has already been used with different methodologies. Kassianov et al. (2005) propose a method to find the best alignment for the two sky images by assuming a single cloud layer in view, aligning the two sky images, and then analytically retrieving the cloud height from the distance in pixels between the centers of the two sky images. In the approach proposed by Nguyen and Kleissl (2014), the authors project the two images onto sky grids, defined by latitude and longitude, with predetermined heights and find the height that aligns the two images in the best way. Both methods are based on the correlation between both images, which implies assuming only one cloud layer. Nguyen and Kleissl (2014) also propose a method to determine the different cloud heights using 2D correlation on the epipolar lines projected on the complementary stereo image, but the correlation must be performed to find each point of interest. Crispel and Roberts (2018) measure the cloud heights but flatten and then rectify the images of the fisheye cameras to obtain a simple stereo system of cameras like the type of perspective camera.

There are studies applying the box voxel concept with two and four cameras using a ceilometer as a validation tool (Kuhn et al., 2018; Nouri et al., 2019). It is worth noting that these models only apply undistortion to the images and

calculate the displacement between the two images, on both axes, and based on this displacement, the height of the points of interest.

In this work, a stereo computer vision model with fisheye cameras is proposed to determine the height of clouds or very distant objects, within a wide field of view, on heuristics of easy physical implementation and low computational complexity. To address these, two premises were considered: i) eliminate the image rectification procedure, which normally requires the transformation of the distorted circular images of the fisheye cameras into undistorted images; ii) eliminate the projection of epipolar lines, even if on the original images.

2. MATERIAL AND METHODS

In this section, the method for camera calibration and the stereo vision model with fisheye cameras used in this work will be discussed. It will also be presented the physical assembly assumed for the experimental system and the specifications of the main instruments that make up the entire experiment.

2.1 Camera model and calibration cameras

In the stereo system, two Vivotek FE9381-EHV cameras are used. Even though they are the same model, they have different distortions, therefore it is initially necessary to perform the calibration of the fisheye cameras separately. Unlike projective (pinhole) cameras, fisheye cameras have a curved lens so their focal length will change with the radius, according to the projective function $f(r)$ and r is related to the image pixel coordinates $\left(r(x,y) = \sqrt{x^2 + y^2}\right)$, Figure 1 shows these relationships. The projective function has polynomial form $f(r) = a_0 + a_1 r + \cdots + a_n r^n$ and the algorithm developed by Scaramuzza et al. (2006) is used to find its coefficients. This omnidirectional camera model allows to approximate the different lenses that compose the fisheye camera to a single lens with its curvature defined by $f(r)$, for which coefficients are estimated from several pictures of a chessboard in several positions regarding the camera. This model is available for use with Matlab (OcamCalib toolbox).

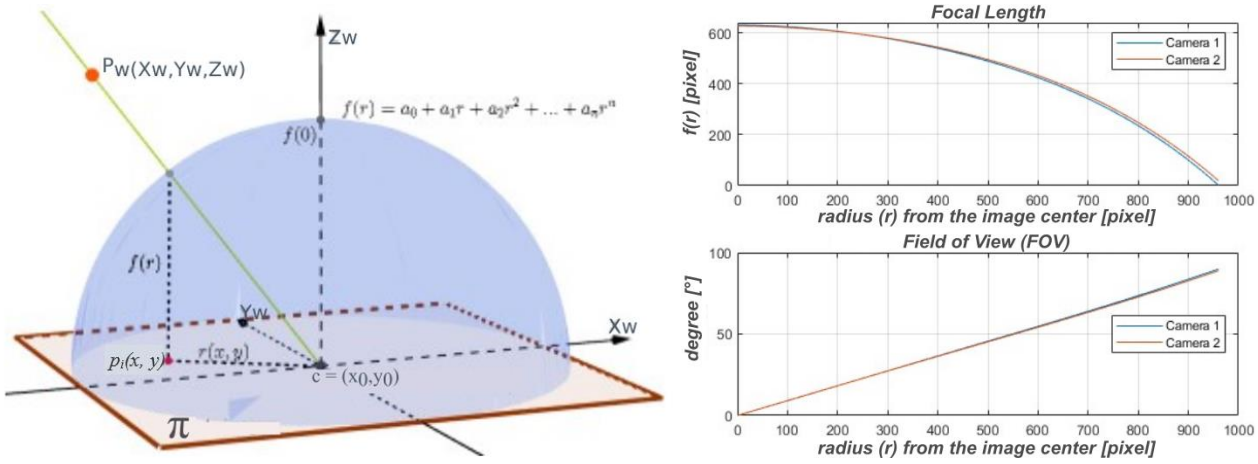

Figure 1. Scaramuzza model fisheye camera and calibration results for two cameras.

The coefficients of the polynomial function $f(r) = a_0 + a_1 r + \cdots + a_n r^n$, obtained in the calibration process are exposes in the Table *1*, $c_{1x} = x_{10}$ and $c_{1y} = y_{10}$, referring to center of the camera 1, and $c_{2x} = x_{20}$ and $c_{2y} = y_{20}$ referring to camera 2. The coefficients a_{11} and a_{21} are zero because $\frac{df(r)}{dr}|_{r=0} = 0$. The first subindex of each of the coefficients refers to camera 1 or 2 respectively.

Table 1: The coefficients of the polynomial function.

Camera 1		Camera 2	
$\boldsymbol{a_{10}}$	630,939	a_{20}	627,381
$\boldsymbol{a_{11}}$	0	a_{21}	0
$\boldsymbol{a_{12}}$	$-683{,}907 \cdot 10^{-6}$	a_{22}	$-514{,}267 \cdot 10^{-6}$
$\boldsymbol{a_{13}}$	$490{,}328 \cdot 10^{-9}$	a_{23}	$996{,}681 \cdot 10^{-10}$
$\boldsymbol{a_{14}}$	$-509.5 \cdot 10^{-12}$	a_{24}	$-240{,}4 \cdot 10^{-12}$
$\boldsymbol{c_{1x} = x_{10}}$	960,09 $[pixel]$	$c_{2x} = x_{20}$	960,27 $[pixel]$
$\boldsymbol{c_{1y} = y_{10}}$	960,31 $[pixel]$	$c_{2y} = y_{20}$	960,07 $[pixel]$

2.2 Fisheye cameras stereo computer vision model

Once both cameras are calibrated, they are positioned such that the x-axis of both cameras are collinear and the optical axes (z), also of both cameras, parallel pointing to the sky aligned with the zenith. In Figure 2, $P_w = [X_w, Y_w, Z_w]$, defines a point in 3D space, world coordinate system (WCS) $[X_w, Y_w, Z_w]$, which is restricted to camera 1 and therefore coincides with its camera coordinate system (CCS). This constraint allows passing in a direct way from WCS to CCS. The image plane of the cameras is defined by the pixel coordinate system (PCS) $[x_{1i}, y_{1i}]$ and $[x_{2i}, y_{2i}]$ for camera 1 and 2 respectively, The PCS may be alternated or referenced as image coordinate system SCI for simplicity or convenience. The image centers $c_1 = [x_{10}, y_{10}]$ and $c_2 = [x_{20}, y_{20}]$ are the optical centers (in pixel) obtained in the previous calibration, as are the projection functions of the cameras as a function of radius $(f_1(r_1), f_2(r_2))$ which are assumed to be the variable focal lengths. r_1 and r_2 are functions of p_{1i} and p_{2i} respectively.

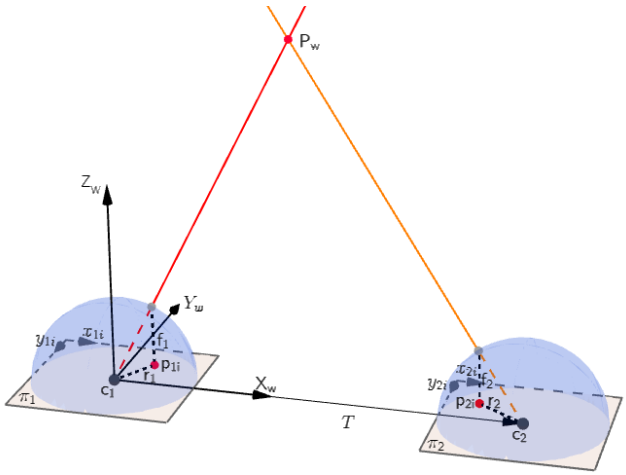

Figure 2: Fisheye stereo vision proposed.

The problem of finding the height of the point P_w of the object of interest is solved by triangulation between the coplanar vectors $\overline{c_1 P_w}$, $\overline{c_2 P_w}$ and $\overline{c_1 c_2} = T$, where T represents the translation vector of camera 2 regarding camera 1. In simple stereo vision systems, as it results in most cases after eliminating the distortion of the images and rectifying them, or in systems where both cameras come aligned and calibrated as a simple stereo system, the measurement of the depth or distance of the objects is obtained from the relation established in the equation resulting in,

$$d(x,y) = \left| [x_2'' - x_1'', \overbrace{y_2'' - y_1''}^{0}] \right| = f\frac{|T|}{z}, \tag{1}$$

where $x_1''$and x_2'' are the pixel coordinates of the camera images after correcting for fisheye lens distortion and after both images are rectified, $|T|$ is the translation vector between the cameras and $z = h$ is the height of point P_W in space. To be able to determine the height in distance units, the pixel size of the camera sensor must be known.

The aim of eliminating the undistortion step and the rectification step is to reduce the error produced in the images, as well as the loss of field of view that the fisheye camera provides. However, it is possible to obtain a relatively large field of view in the undistortion process by making the chosen f smaller than the $f(0)$ provided by the manufacturer. Defining the field of view (W) a new f will be obtained from the relation $\tan W = \frac{NP}{2f}$ where NP is the image size in pixels. However, this process compresses the image in the center and dilates it at the edges (see Figure 3), introducing even more errors and distortions due to the reverse transformations performed from the function $f(r)$ obtained in the calibration of both cameras. One should also consider the change of the scale factor in the case of using a different focal length than the original camera. Figure 3 summarizes the above.

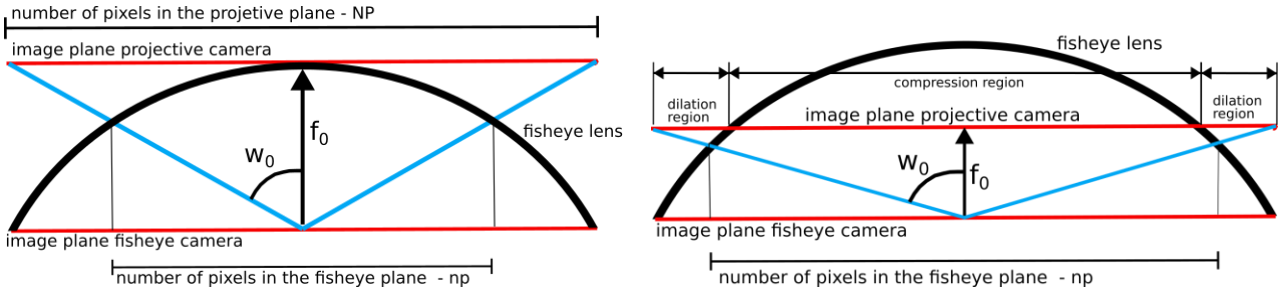

Figure 3: Undistortion fisheye camera image.

To avoid these problems and reduce the processing required in these steps, the omnidirectional camera model proposed by Scaramuzza is used, which allows performing a transformation from 2D to 3D space, establishing the relationship between a given point in the image and a 3D vector that originates from the center of the image. The proposed omnidirectional camera model treats the system as a single compact system, eliminating all the intermediate lenses that could exist in the cameras. Thus, in this work a simpler fisheye stereo vision system is proposed, in which Eq. (1) is not utilized, mainly because variable $f(r)$ is assumed.Another problem that must be considered in triangulation is the case in which the vectors, due to difficulty in aligning distant cameras, $\overline{c_1P_w}$ and $\overline{c_2P_w}$ are not coplanar, and consequently do not intersect. This is also the result of the large number of variables implicit in the optical model (intrinsic parameters) used to model fisheye cameras; the extrinsic parameters that bring the cameras closer to an optimal alignment; the cameras' own errors. Figure 4 shows graphically the problem of non-coplanarity between the vectors. A possible way to solve this problem is to find the smallest segment between both lines, equivalent to finding the orthogonal segment between them.

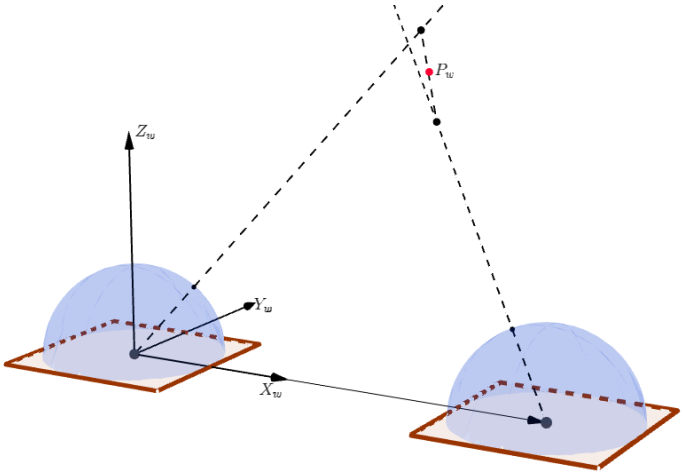

Figure 4: Fisheye stereo vision problem of non-coplanarity of vectors.

The camera model proposed by Scaramuzza is based on the following hypothesis,

$$P_w = \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} = \lambda \begin{bmatrix} x_i - c_x \\ y_i - c_y \\ f(r) \end{bmatrix}, \tag{2}$$

this model treats P_w as a vector in $3D$ space, so this vector is collinear to the vector defined by $\left[x_i - c_x, y_i - c_y, f(r)\right]^T$, where $r = \sqrt{(x_i - c_x)^2 + (y_i - c_y)^2}$ and λ is a scale factor that will allow the points to be mapped onto the line defined by the vector $\overline{cP_w}$ (see Figure 1). It is therefore possible to extend this single omnidirectional camera model to a stereo model (see Figure 2). Assuming the constraint that the coordinate system of camera 1 (in this case CCS) is coincident with the WCS, the following assumption holds,

$$P_w = \overbrace{\lambda_1 \underbrace{\begin{bmatrix} x_{1i} - c_{1x} \\ y_{1i} - c_{1y} \\ f(r_1) \end{bmatrix}}_{v_1}}^{Cam\ 1} = \overbrace{\lambda_2 [R] \underbrace{\begin{bmatrix} x_{2i} - c_{2x} \\ y_{2i} - c_{2y} \\ f(r_2) \end{bmatrix}}_{v_2} + [T]^T}^{Cam\ 2}, \tag{3}$$

where $c_{1x}, c_{1y,}\, c_{2x}$ and c_{1y} correspond to the coordinates of the image center $c = (c_x, c_y)$ of both cameras respectively, x_{1i}, y_{1i}, x_{2i} and y_{2i} are the pixel coordinates of the image (PCS, the position 0,0 corresponds to the upper-left corner of the image), $[R]$ and $[T]$ are the rotation matrix and the translation vector, respectively, of camera 2 regarding camera 1. However, $[R]$ is disregarded since it is assumed that the cameras are aligned and leveled regarding their coordinate axes, where $[T] = \frac{\mathrm{T}\,[mm]}{pixel\ size\,[mm]}$ must be in pixel unit. To make the system of equations more compact, everything is expressed in image coordinates,

$$\lambda_1 \overbrace{\begin{bmatrix} x_{1c} \\ y_{1c} \\ f(r_1) \end{bmatrix}}^{v_1} = \lambda_2 \overbrace{\begin{bmatrix} x_{2c} \\ y_{2c} \\ f(r_2) \end{bmatrix}}^{v_2} + \overbrace{\begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}}^{T}, \tag{4}$$

reorganizing we obtain,

$$\overbrace{\begin{bmatrix} x_{1c} & -x_{2c} & -t_x \\ y_{1c} & -y_{2c} & -t_y \\ f(r_1) & -f(r_2) & -t_z \end{bmatrix}}^{A} \overbrace{\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ 1 \end{bmatrix}}^{\lambda} = \overbrace{\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}}^{b}, \tag{5}$$

thus, it results in a system of inhomogeneous linear equations over determined, and as presented in Figure 4 the lines do not necessarily intersect and therefore the problem can be solved as a minimization problem, transforming it into a system of homogeneous equations, according to Eq. (5). Knowing that b will only take the value 0 in the ideal case of the system and for the trivial solution $\lambda = 0$ (which is not of interest). This type of system is easy to solve using singular value decomposition (SVD). An in-depth description of the subject can be found in Hartley and Zisserman (2004, Appendix 5). Thus, given a matrix A, which meets the condition of numbers of rows greater than or equal to the number of columns, the objective is to find λ that minimizes b, with λ restricted to ‖λ‖=1, applying SVD to A one obtains $SVD(A_{mxn}) = U_{m\times n} D_{n\times n} V^T_{n\times n}$ and the values of λ are defined by the values of the last column of V, as follows,

$$\lambda = \begin{bmatrix} \frac{v_{13}}{v_{33}} \\ \frac{v_{23}}{v_{33}} \\ \frac{v_{33}}{1} \end{bmatrix}, com\ \lambda_1 = \frac{v_{13}}{v_{33}}\ e\ \lambda_2 = \frac{v_{23}}{v_{33}}, \tag{6}$$

2.3 Physical installation and materials

The two cameras were installed on the outside of the roof of the Energy and Environment Laboratory (LEA) at UnB (see Figure 5a). The specifications of both cameras are detailed in Table 2. For the connection between the cameras, a wireless connection system consisting of two antennas and an Ethernet switch was used to have a single connection point to the cameras (see Figure 5b,c). Tripods and a base for leveling and aligning the cameras were used to support the cameras (see Figure 5e,d). The leveling was performed by a digital inclinometer (with 0.10° resolution). The distance between the cameras was checked by a digital rangefinder (with 0.10 m resolution) and tape measure.

(a) (b) (c) (d) (e)

Figure 5: (a) Rooftop physical installation of the cameras. (b) and (c) Antennas for communication with hub. (d) Camera fisheye with with base, support, and connection box. (e) Tripod and a base for leveling and aligning the cameras.

Table 2: Camera specifications.

VIVOTEK FE9381-EHV Technical Specifications	
Pixel size	2.4 μm x 2.4 μm
Maximum resolution	1920 x 1920 pixels
Focal length	1.47 mm
Field of view	180° horizontal/vertical/diagonal

Finally, with the system installed and leveled, it is necessary to perform the x-axis alignment of the cameras, one regarding the other. The calibration of the rotation between the cameras, on the horizontal plane, is the most difficult to perform manually. On the other hand, it is simple through post-processing, using the images of the sky with the Sun as reference. As can be seen in Figure 6, a rotation of 5.6° counterclockwise was applied to the image obtained by camera 2.

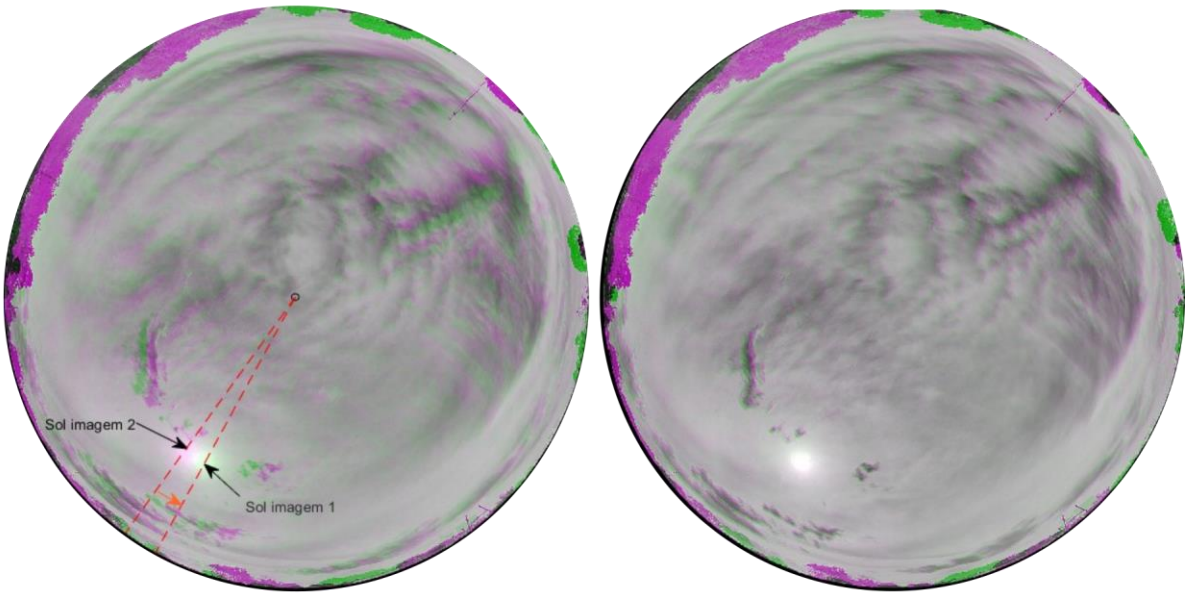

Figure 6: Calibration of the rotation between the cameras.

3. RESULTS AND DISCUSSIONS

The validation of the proposed stereo vision model with fisheye cameras, to be applied in a cloud height measurement system, is performed using a drone as a reference point (P_w) with known height. The drone allows validating the height measurement in a wide field of view. For the experiment, the drone was positioned at desired heights, at each of these positions synchronized photos were taken with the drone camera and with both cameras of the stereo system pointed to the sky. The experiments were carried out at night to facilitate the identification of the drone in the images because it has a white light at its bottom. The photos obtained with the drone camera have the height information and geographic coordinates related to its positioning.

The pictures were taken by sectors (right), identified with numbers 1 to 5, and at different heights identified with colors (Figure 7 left).

Figure 7: Sectors or quadrants – left. Sequence of pictures at 490 m (magenta), 390 m (yellow), 290 m (blue) and 190 m (red) – right.

The drone used was DJI, model Mavic Air 2. Its maximum flight altitude is 500 m, and an accuracy in stationary flight of ±0.5 m vertical and ±1.5 m horizontal (with GPS positioning). In its lower part, it has led lighting, which allows a location in night flights. It has a stabilization system with a maximum control speed of 100 °/s with an angular vibration range of ±0.01°. Therefore, this model provides an extremely reliable reference point in space for the measurement range used 90 m - 490 m. The relative height of the cameras to the drone take-off origin point is 10 m.

Figure 8: DJI Mavic Air 2 drone, used for validation.

The validation of the model with cloud images is usually performed with a ceilometer as a reference instrument. However, the ceilometer only makes it possible to measure the height of the clouds passing over its sensor. Therefore, the use of the drone allows a validation of the stereo vision system on a larger horizontal range.

Figure 9 shows the acquisition of the positioning (in pixels) of the drone on the camera images, which is manually obtained.

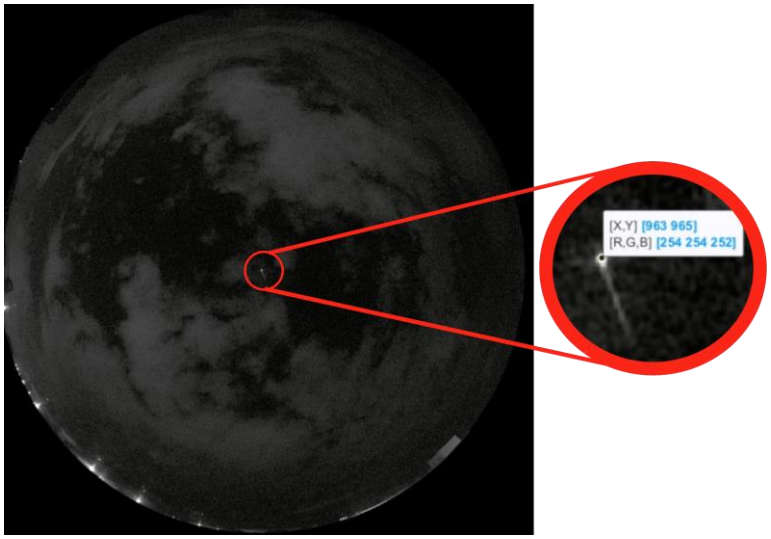

Figure 9: Acquisition of the positioning (in pixels) of the drone in the images.

Analyzing the errors represented in the left graph of Figure 10, it is observed that there is an overestimation bias of the drone height in the positions indicated by picture 2 and 3 at almost all heights, the exception being the drone position indicated by photo 2 at 490 m. It is also observed a higher accuracy and precision in the measurements of the positions indicated by picture 5, which are the ones that have greater proximity to the center of the computer vision system. The absolute percentage errors relative to each measurement, as shown in the right graph of Figure 10, show that the accuracy and precision of the measurement increases with increasing height, as the error is smaller and more uniform. The global metrics used for system evaluation were Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE) and Root Mean Square Error (RMSE). For the calculation of these metrics the 20 drone positions were considered and the results for both cameras are shown in Table 3.

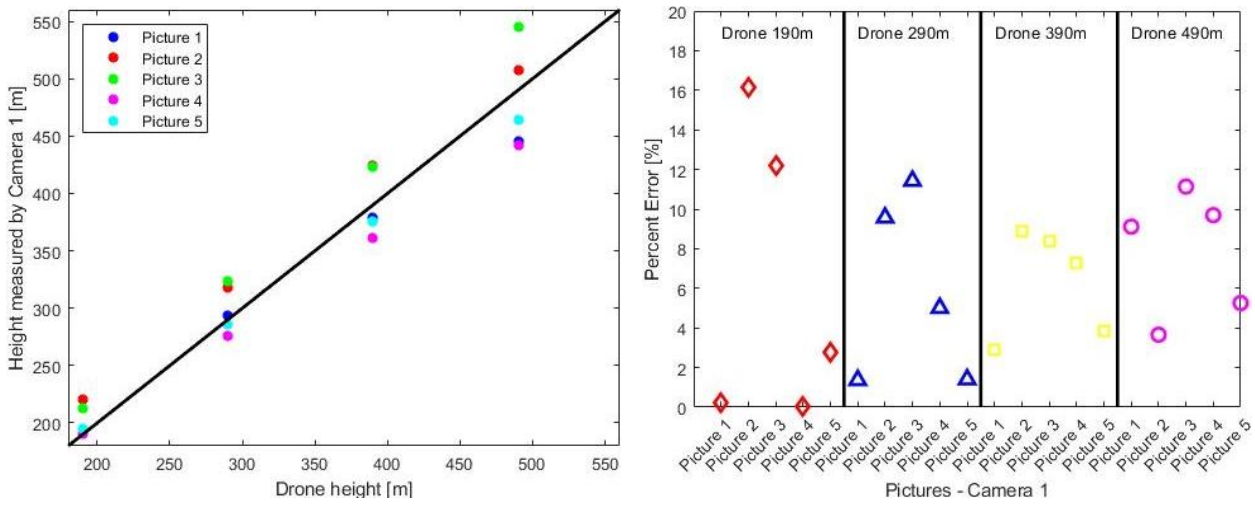

Figure 10: Absolute Error (right) and percent error (left).

Table 3: Global metrics considering the twenty drone positions for the two cameras.

ERROR	Camera 1	Camera 2
MAE [m]	22,78	21,53
MAPE [%]	6,52	7,34
RMSE [m]	7,86	9,70

The first observation to infer is that the error is related to the position radius of the drone in the images, but no such relationship was found when determining the drone positioning radius in each picture. As a conclusion of this experimental phase, it is inferred that the cameras should be aligned and leveled again using some equipment that allows

to obtain a substantial improvement, especially about leveling in the horizontal plane. To verify this hypothesis, the system was readjusted, and new measurements were taken, this time only considering the highest possible height of 490 m and a height of 90 m, both in relation to the cameras, as shown in Figure 11.

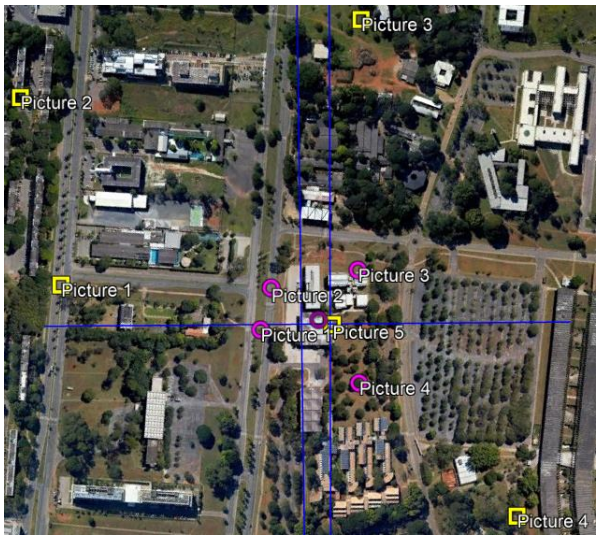

Figure 11: Sequence of pictures at 490 m (magenta), 90 m (yellow) after recalibration.

The values of the heights measured by camera 1 are presented in the left graph of Figure 12, where it is possible to observe that there is an improvement in the measurements. This is also verified through the absolute percentage errors of each of the measurements, which can be observed in the right graph of Figure 12. The highest absolute percentage error is 11.09% relative to camera 1 for the height of 90 m and 6.86% for the height of 490 m, also referring to camera 1. The results indicate a substantial improvement after the physical recalibration of the system.

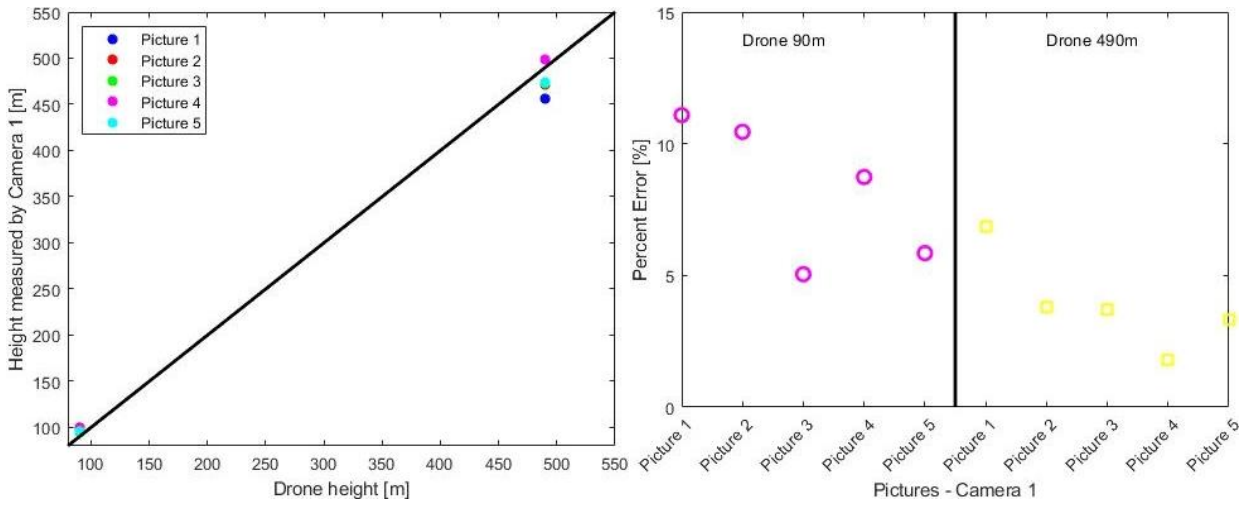

Figure 12: Absolute Error (right) and percent error (left) after recalibration.

The global system metrics for each of the cameras are shown in Table 3 and confirm the improvement in measurement accuracy and precision.

Table 4: Global metrics considering the twenty drone positions for the two cameras after recalibration.

ERROR	Camera 1	Camera 2
MAE [m]	13,21	13,95
MAPE [%]	6,06	5,87
RMSE [m]	6,76	6,55

4. CONCLUSION

The proposed model and the new validation method using a drone brought satisfactory results. Therefore, they proved to be suitable for use in the implementation, configuration and adjustment of a system that enables the determination of cloud height to be used in short-term solar forecasting. However, its use can be extended to other applications where this data is relevant, such as for airport operation and for meteorological applications.

It is possible to establish a stereo vision model with fisheye cameras to be implemented in the measurement of objects at relatively high distances with results lower than 10% without the need for undistortion and rectification of the images, which would decrease the processing time.

Finally, for system improvements, it is proposed to employ a base for each of the cameras in the system, in which it has mechanical systems that enable better calibration of leveling and positioning, in addition to providing greater robustness against unevenness caused by external forces.

5. ACKNOWLEDGEMENTS

This research was funded by the resources made available by DPG Public Notice No. 0010/2023 0010/2023 "Apoio à execução de projetos de pesquisas científicas, tecnológicas e deinovação de discentes de pós-graduação" and partially by FAP-DF through the research project entitled "Estudo de Sistema Solar Híbrido Fotovoltaico Térmico com Concentrador Solar e Absorvedor de Calor" (project number 00193.0000226/2019-72).

6. REFERENCES

Abraham, S., & Förstner, W. (2005). Fish-eye-stereo calibration and epipolar rectification. *ISPRS Journal of Photogrammetry and Remote Sensing*, *59*(5), 278–288. https://doi.org/10.1016/j.isprsjprs.2005.03.001

Cai, C., & Qiao, R. (2021). An Improved Calibration Method for Stereo Fisheye Vision. *Proceedings of the 33rd Chinese Control and Decision Conference, CCDC 2021*, 3158–3162. https://doi.org/10.1109/CCDC52312.2021.9602138.

Cao, Y., Wang, H., Zhao, H., & Yang, X. (2022). Neural-Network-Based Model-Free Calibration Method for Stereo Fisheye Camera. *Frontiers in Bioengineering and Biotechnology*, *10*. https://doi.org/10.3389/fbioe.2022.955233

Crispel, P., & Roberts, G. (2018). All-sky photogrammetry techniques to georeference a cloud field. *Atmospheric Measurement Techniques*, *11*(1), 593–609. https://doi.org/10.5194/amt-11-593-2018.

Hartley, R.; Zisserman, A. Multiple View Geometry in Computer Vision, (2004). Second Edi ed. Cambridge University Press.

Kassianov, E., Long, C. N., & Christy, J. (2005). Cloud-Base-Height Estimation from Paired Ground-Based Hemispherical Observations. *JOURNAL OF APPLIED METEOROLOGY*, *44*, 1221–1233.

Khomutenko, B., Garcia, G., & Martinet, P. (2017, January 31). Direct fisheye stereo correspondence using enhanced unified camera model and semi-global matching algorithm. *2016 14th International Conference on Control, Automation, Robotics and Vision, ICARCV 2016*. https://doi.org/10.1109/ICARCV.2016.7838761

Kuhn, P., Wirtz, M., Killius, N., Wilbert, S., Bosch, J. L., Hanrieder, N., Nouri, B., Kleissl, J., Ramirez, L., Schroedter-Homscheidt, M., Heinemann, D., Kazantzidis, A., Blanc, P., & Pitz-Paal, R. (2018). Benchmarking three low-cost, low-maintenance cloud height measurement systems and ECMWF cloud heights against a ceilometer. *Solar Energy*, *168*, 140–152. https://doi.org/10.1016/j.solener.2018.02.050.

Nguyen, D., & Kleissl, J. (2014). Stereographic methods for cloud base height determination using two sky imagers. *Solar Energy*, *107*, 495–509. https://doi.org/10.1016/j.solener.2014.05.005.

Nouri, B., Kuhn, P., Wilbert, S., Hanrieder, N., Prahl, C., Zarzalejo, L., Kazantzidis, A., Blanc, P., & Pitz-Paal, R. (2019). Cloud height and tracking accuracy of three all sky imager systems for individual clouds. *Solar Energy*, *177*, 213–228. https://doi.org/10.1016/j.solener.2018.10.079.

Scaramuzza, D., Martinelli, A., & Siegwart, R. (2006). A toolbox for easily calibrating omnidirectional cameras. *IEEE International Conference on Intelligent Robots and Systems*, 5695–5701. https://doi.org/10.1109/IROS.2006.282372.

Togawa, A. (2014). Previsão de irradiância solar por rastreamento de nuvens utilizando imagens. Universidade de Brasília (UNB).

Trucco, E.; Verri, A. (1998). Introductory techniques for 3-D computer vision. Englewood Cliffs: Prentice Hall.

7. RESPONSIBILITY NOTICE

The authors are the only responsible for the printed material included in this paper.