

USE OF FISHEYE PARROT BEBOP 2 IMAGES FOR 3D MODELLING USING COMMERCIAL PHOTOGRAMMETRIC SOFTWARE

D. Pagliari^{1,*}, L. Pinto¹

¹ Politecnico di Milano, DICA - Geodesy and Geomatics section, Piazza Leonardo da Vinci 32, 20133 Milan, Italy
(diana.pagliari, livio.pinto)@polimi.it

Commission II, WG I

KEY WORDS: Fisheye, Distortions, Camera Calibration, Parrot Bebop 2, 3D modelling, Agisoft PhotoScan, Pix4D Mapper

ABSTRACT:

Fisheye camera installed on-board mass market UAS are becoming very popular and it is more and more frequent the use of such platforms for photogrammetric purposes. The interest of wide-angles images for 3D modelling is confirmed by the introduction of fisheye models in several commercial software packages. The paper exploits the different mathematical models implemented in the most famous commercial photogrammetric software packages, highlighting the different processing pipelines and analysing the achievable results in terms of checkpoint residuals, as well as the quality of the delivered 3D point clouds. A two-step approach based on the creation of undistorted images has been tested too. An experimental test has been carried out using a Parrot Bebop 2 UAS by performing a flight over an historical complex located near Piacenza (Northern Italy), which is characterized by the simultaneous presence of horizontal, vertical and oblique surfaces. Different flight configurations have been tested to evaluate the potentiality and possible drawbacks of the previously mentioned UAS platform. Results confirmed that the fisheye images acquired with the Parrot Bebop 2 are suitable for 3D modelling, ensuring accuracies of the photogrammetric blocks of the order of the GSD (about 0.05 m normal to the optic axis in case of a flight height equal to 35 m). The generated point clouds have been compared to a reference scan, acquired by means of a MS60 MultiStation, resulting in differences below 0.05 in all directions.

1. INTRODUCTION

The use of fisheye camera has becoming very popular because of their great advantages in terms of field of view, if compared to the one of rectilinear cameras, coupled with their low-cost market availability. Firstly, these systems were commercialized for entertainment purposes (e.g. photography, realization of personal or commercial videos), but recently they have been installed on several Unmanned Aerial Systems (UAS), which are more and more often used for photogrammetric purposes too. In fact, UAS have gained a lot of popularity, in recent years, thanks to the progress on both control and navigation sensors, together with the development of high quality and lightweight cameras. Indeed, the combination of Global Navigation Satellite System (GNSS) and low-cost inertial systems allows flying by maintaining a correct balance of the UAS platform while it is moving. Due to the presence of high quality imaging sensor, coupled with a huge development of dedicated commercial software packages, the data acquired with UAS could be very useful for geometric surveys of both infrastructure and terrain. The main advantage connected to the use of such platforms is the possibility of acquiring images by following precise flight path, guaranteeing a sufficient overlapping and a stable image scale, which are all necessary conditions for photogrammetric processing. The view from above enables a global vision of the investigated object, overcoming the occlusion problems typical of ground acquisitions performed using classical topographic methods (Eisenbeiss, 2004, Remondino et al 2011, Gagliolo et al 2017). In many UAS, autopilot systems that guarantee a high level of flight automation are embedded. Moreover, the last generation of drones has often a collision avoidance system, thus guarantying a minimum safety distance between the investigated object and the flying platform during the survey (Accardo et al., 2013). The

use of such systems coupled with fisheye lenses is very interesting in case of surveying of buildings and infrastructures. The Field of View (FOV) of fisheye lenses is very wide (about 180°), resulting in the use of projection models that are different from the classical pinhole model commonly used in case of rectilinear lenses (Kannala and Brandt, 2006, Perfetti et al., 2017). In fact, when the angle of the incoming ray tends to be equal to 90° the classical projection model tends to a singularity point (infinity). The most common projection model used for fisheye lenses is based on the equidistant one; however, several theoretical mapping functions can be used to describe the relationship between the incoming ray and the object point (see Kannala and Brandt, 2006). In order to produce accurate 3D models, it is important to consider also the residual distortions describing the discrepancies between real lenses and the chosen mapping function, in addition to the theoretical model. Furthermore, the non-linear variability of the Ground Sample Distance (GSD) of fisheye lenses could represent an obstacle during the matching phase for automatic interest operators, influencing the quality of the estimated external orientation parameters and the delivered photogrammetric products. On the other hand, fisheye lenses represent a great opportunity because they can significantly reduce the amount of requested images, guarantying at the same time acquisition geometries that are interesting from a pure photogrammetric point of view (Bolognesi et al., 2015, Covas et al., 2015). The Parrot Bebop 2 (from now on Bebop) combines a number of the drone characteristics discussed so far. It is a small quadcopter equipped with a 14 Megapixel fisheye camera, with a fixed inclination of 30° (looking downward). Together with the DJI Phantom, it is considered a commercial UAS useful for both professional aerial acquisitions and for entrainment purposes

* Corresponding author

(e.g. video and photographic services) (de Miguel Molina and Ona, 2018).

The presence of an oblique camera on-board represents a challenge for classical photogrammetric software packages, but it is very interesting because it allows overcoming the occlusion problems that are typical of nadir flights. This task has been addressed by the Italian Society of Photogrammetry and Topography (SIFET) dedicating a benchmark to analyze the potential and critical aspects of the use of UAS oblique images. (Piras et al., 2017).

The Bebop installs on-board a number of different sensors helpful for stabilizing its flight. A pressure sensor measures the air pressure and analyzes the flight altitude beyond 4.9 m. These data are completed with information acquired with an ultrasound sensor up to that height. The Bebop navigation systems is composed by several sensors, which data are analyzed and processed by the onboard computer. A GNSS (GPS+GLONASS+GALILEO) chipset is used for the UAS geolocalization, but also for speed measurements and flight height stabilization. A 3-axis accelerometer is used for measuring the drone position and its linear speed. In addition, the information acquired with a 3-axis magnetometer concurs in the definition of the platform position. The attitude of the UAS is measured by means of a 3-axis gyroscope. Because of its low weight, the Bebop can guarantee a flight autonomy up to 25 minutes. Furthermore, it can be remotely controlled via smartphone/tablet (via Wi-Fi connection). The low-cost (about € 600), its easiness of use and the good quality of the optic make it one of the best ready-to-go solution available on the market for low height aerial image acquisition. In this paper, an experimental test carried out with a Bebop is presented and discussed. The aim of this study is to evaluate the applicability of such UAS for photogrammetric purpose, evaluating the quality of the resultant 3D model in terms of metric accuracy and completeness.

Few examples of the use of Bebop UAS for navigation and mapping are present in literature (see among the others Huang et al., 2017, Yan et al., 2017, Taufik et al., 2016). However, to best of our knowledge, the authors use images extracted from the videos or from manual shots; these images are characterized by a 1920x1080 resolution (cropped format) and the radial distortions are minimal, if compared to the ones of the images acquired with the higher resolution, which can be assimilated to spherical fisheye images.

Different software packages and approaches have been tested to highlight the impact of different fisheye projection functions and distortion models on the resultant photogrammetric products.

The rest of the paper is organized as follows. In Section 2, the different mathematical formulation of fisheye projection models embedded in the most common commercial software are introduced. In Section 3, the experimental test carried out using the Bebop UAS is presented and discussed, by comparing the processing pipeline used for the different tested software packages. The obtained results have been evaluated in terms of residuals of the bundle block adjustment, quality of the photogrammetric products and by comparing the generated point clouds with a reference scan. Finally, few conclusions are drawn in Section 4.

2. FISHEYE PROJECTION MODELS

Fisheye cameras are characterized by a wide FOV and a short focal length; however, these two parameters are not sufficient to distinguish between rectilinear and fisheye cameras (Perfetti et al., 2017). Because of the peculiarity of fisheye lenses, a mapping function different from the classical pinhole model used for

rectilinear images is required. In fact, those models tend to infinity when the angle of the incoming ray tends to 90°, thus resulting in the impossibility of a FOV of 180° using a perspective projection. Instead, for fisheye lenses a FOV of 180° is always possible, meaning that a different optical projection is needed. It is also important to point out the difference between the mapping function and the distortions, due to a non-perfect correspondence between the physical lenses and the chosen mapping function. Fisheye camera models have been studied from decades in Photogrammetry and several models have been formalized (see for instance Steffen and Förstner, 2005 Schwalbe, 2005, Schneider et al., 2009, Barazzetti et al, 2017). Typically, the residual distortions are split between two main components, namely radial and tangential; the latter is often neglected because its effect is minimal when compared to the one of the radial distortion.

The use of wide-angle cameras is an important task also in Computer Vision because of the great advantages that this kind of images could represent for many robotic vision tasks (see Courbon et al., 2007, Hughes et al., 2010). According to Courbon et al. (2007), the fisheye camera models can be classified in three main groups: 1) The camera geometry is based on the conventional pinhole model, which is then corrected adding radial distortions. The fisheye images are transformed to follow the camera model, but their applicability is limited and could easily lead to huge approximations. 2) The mapping between the image points and the 3D points is done considering the incidence angle between the image point and the ray direction, with respect to the image centre. A review of the main type of optical projection is given in Kannala and Brandt (2006). 3) A generic model used for catadioptric and/or omnidirectional cameras is adapted to fisheye case. The relationship between image and object points is described by a polynomial function.

Recently fisheye camera models have been introduced in the most famous and commonly used photogrammetric software packages, such as Agisoft PhtoScan – version 1.3.4 (PS) and Pix4D Mapper - version 4.0.25 - (P4), and a dedicated camera calibration tool has been embedded in Matlab 2017b (MAT). However, the different software exploits different mathematical models.

The PS fisheye projection model is based on an equidistance projection, while the distortions are described using an extended version of Brown's model (Brown, 1971), modelling both radial and tangential distortions combined with affinity and shear parameters (El-Hakim, 1986). This could be very important especially in case of images acquired from moving platforms, such as UASs. The PS model is described by:

$$\begin{aligned} x_d &= \frac{f \cdot X}{\sqrt{X^2 + Y^2}} \arctg\left(\frac{\sqrt{X^2 + Y^2}}{Z}\right) + c_x + \Delta x_{dist} \\ y_d &= \frac{f \cdot Y}{\sqrt{X^2 + Y^2}} \arctg\left(\frac{\sqrt{X^2 + Y^2}}{Z}\right) + c_y + \Delta y_{dist} \end{aligned} \quad (1)$$

where f is the focal length, X, Y, Z are the object coordinates of a generic point in a camera centered reference system, c_x, c_y are the coordinates of the principal point and $(\Delta x_{dist}, \Delta y_{dist})$ take into account of radial, tangential distortions as well as affinity and shear parameters.

P4 uses a projection model that maps the angle θ between the incident ray and the camera, which is defined as:

$$\theta = \frac{2}{\pi} \arctg\left(\frac{\sqrt{X^2 + Y^2}}{Z}\right) \quad \text{with } \theta \in [0,1] \quad (2)$$

The relationship between object and image coordinates is modelled as:

$$\begin{pmatrix} x_d \\ y_d \end{pmatrix} = \begin{bmatrix} C & D \\ E & F \end{bmatrix} \begin{pmatrix} \rho \cdot X / \sqrt{X^2 + Y^2} \\ \rho \cdot Y / \sqrt{X^2 + Y^2} \end{pmatrix} + \begin{pmatrix} c_x \\ c_y \end{pmatrix} \quad (3)$$

where $\rho = \theta + p_2\theta^2 + p_3\theta^3 + p_4\theta^4$, p_2, p_3, p_4 are the coefficients of a polynomial function, C, D, E, F are the coefficients that allow to map the undistort image coordinates into the distorted ones (x_d, y_d) . The diagonal element of this matrix can be related with the focal length, considering that:

$$f = \frac{2C}{\pi} \quad (4)$$

In case of Bebop camera the parameters D and E are set equal to zero, while C is imposed to be equal to F. P4 allows to compensate for the rolling shutter effect too, however this option is not set as default by the software when the Bebop camera is recognized from the EXIF file, so it has not be used for the work here discussed. Details about how the rolling shutter correction is implemented in P4 can be found in Vautherin et al (2016).

In MAT, a new set of functions that enables to calibrate fisheye cameras and to undistort their images has been embedded. Those functions rely on the general model for calibrating omnidirectional cameras presented in Scaramuzza et al. (2006). It assumes that the function that describes the relationship between an image point and the 3D ray generated by the object point could be represented with a Taylor series expansion. The calibration procedure consists in the estimation of those parameters via 4-step least square minimization, finalizing their estimation with a non-linear solution. The general model presented by the authors is

$$\begin{pmatrix} X_c \\ Y_c \\ Z_c \end{pmatrix} = \lambda \begin{pmatrix} x \\ y \\ a_0 + a_2\rho^2 + a_3\rho^3 + a_4\rho^4 \end{pmatrix} \quad (5)$$

where X_c, Y_c, Z_c are the object coordinates of the camera projection center, λ is a scalar factor, x, y are the ideal projections of the 3D point, a_0, a_1, a_2, a_3 are the polynomial coefficients to be estimated during the calibration procedure and ρ is equal to $\sqrt{x^2 + y^2}$. The real distorted coordinates (x_d, y_d) are related to the real ones by the following relationship:

$$\begin{pmatrix} x_d \\ y_d \end{pmatrix} = \begin{bmatrix} c & d \\ e & 1 \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} c_x \\ c_y \end{pmatrix} \quad (6)$$

where c, d, e are the coefficients which model the stretching and the distortions, compensating from sensor-to-lens misalignment. The model presented so far is valid for omnidirectional cameras, whose fisheye lenses represent a particular case. For this reason during the calibration procedure c is automatically set equal to one and d and e are set equal to zero, resulting in a stretch matrix equal to the identity. Under this simplification hypothesis the resultant model is:

$$\begin{aligned} x_d &= \frac{X}{Z} (a_0 + a_2\rho^2 + a_3\rho^3 + a_4\rho^4) + c_x \\ y_d &= \frac{Y}{Z} (a_0 + a_2\rho^2 + a_3\rho^3 + a_4\rho^4) + c_y \end{aligned} \quad (7)$$

After the end of the camera calibration procedure, MAT allows to generate undistort images, creating at the same time also the intrinsic parameters of the correspondent virtual frame camera.

The MAT and the P4 models are based on different projecting principles; however, they both use a polynomial function and a stretch matrix for distortion mapping. Because the function used by P4 is normalized in the interval $[0,1]$, it is difficult to directly compare the estimated coefficients values. Nevertheless, it is worth noticing that the value of a_0 in MAT corresponds to the focal length value estimated by P4.

3. EXPERIMENTAL TEST

3.1 The test area

In order to evaluate the applicability of Bebop for photogrammetric purposes and to test the use of oblique fisheye images in commercial close-range photogrammetric software packages, an experimental test has been conducted over the Caratta Castle (located near Piacenza – Northern Italy). The area is located in the countryside and it is approximately 1 ha wide and it is composed by several historical buildings, two courtyards and colonnaded constructions. In the last centuries, they have been used as farmstead. The buildings are quite damaged and could represent a typical application for Architectural drone survey, especially because of the simultaneous presence of horizontal, vertical and oblique surfaces, namely courtyard floors, building façades and roofs (see Fig.1). The historical complex is characterized by the presence of elements with different heights, among which the highest ones are the two towers (16 m above ground level). Considering all these factors, this area was selected for the experimental test for evaluating the potentiality of the Bebop UAS.

Two different grid flights have been realized over the investigated area. They have been projected to guarantee an average GSD below 0.05 m. The Bebop camera nominal focal length is equal to 1403 pixels and the nominal pixel size is equal to 1.4 μm . Because of the use of oblique fisheye images, it is important to point out that there is a huge variability of the GSD throughout the images. For this reason, two different GSDs have been considered to carefully plan the flights. The nadir GSD corresponds to $7 \cdot 10^{-4} \cdot d$, where d represents the distance between the camera and the object, while the GSD normal to the optic axis redoubles because of the camera inclination of 30°. In order to guarantee a GSD normal to the optic axis equal to 0.05 m (and consequently a nadir GSD of about 0.025 cm) the resultant flight height is equal to 35 m. The images have been acquired with the Bebop embedded Sunny 180° fisheye camera (see Table 1 for optics technical details).

Specification	Parameters
sensor	CMOS 14 Mpx
optical	Sunny 180° eye 1/2.3" aperture
lens stabilizer	3-axis digital system
video resolution (pixel)	1920x1080 (30 fps)
image resolution (pixel)	4096 x 3320
nominal focal length (pixel)	1403
nominal pixel size [μm]	1.4

Table 1. Technical specifications of Bebop image acquisition system

A total of 389 images have been acquired over the investigated area. Each one of the two grids flight was organized with 8 strips along E-W direction and 7 strips along N-S direction, with overlapping of about 90% along flight direction and of about 80% along crossed flight direction. The flights were performed in a full-automatic way, controlling the UAS from the tablet, via Pix4D Capture app.

A geodetic network with 5 vertexes (see Fig.1) has been materialized to measure the position of 21 targets (i.e. black and white square targets, with a side of 0.30 m) by means of a Leica MS60 MultiStation (Fagandini et al, 2017).

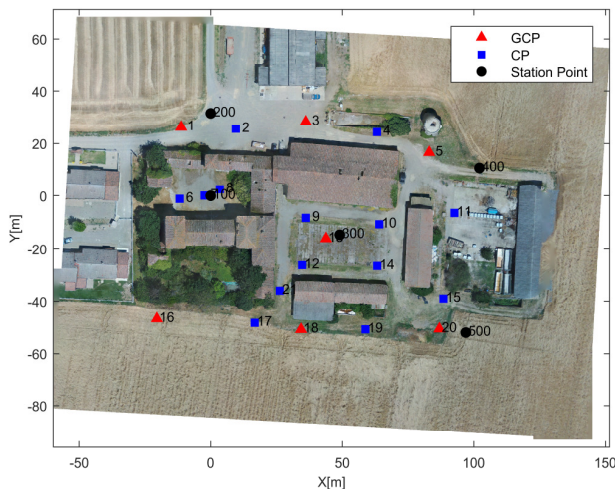


Figure 1. Surveyed area and geodetic network scheme

Among them, 7 targets have been used as Ground Control Points (GCPs) (6 of them located outside the castle and one located in the middle of the main courtyard). The other points have been used as Check Points (CPs), guaranteeing the independence of the measurements. The coordinates of two station points were measured with a GNSS receiver, in order to georeference the network in the ETRF2000(08) global reference frame. Finally, the geodetic network has been transformed in a local reference system, with origin in the point 100, Z-axis along the vertical direction and Y-axis orientated along the direction defined by the points 100 and 200.

3.2 Image Processing

Different software packages and approaches have been tested to understand if it is possible to process fisheye oblique images with the most common commercial photogrammetric software packages. The standard image processing workflow has been followed for P4; it represents the natural solution because Parrot and P4 have been developed with a great effort between the two firms. Instead, two different approaches have been used for PS. A first solution has been computed by following the standard workflow suggested for the processing of a photogrammetric block, using the embedded fisheye camera model, whereas a two-step approach has been used for the MAT generated images. Firstly, they have been converted into frame images using MAT, and then they have been processed in PS to perform the bundle block adjustment. All the processing have been done using a Dell Inspiron 15 7000 Gaming, with 16GB (2400MHz, DDR4), 512GB PCIe Solid State Drive and dedicated NVIDIA GeForce GTX 960M (4GB GDDR5) graphics memory.

For each one of the analysed software, several image block configuration have been considered: 1) grid flight - using all the acquired images, 2) N-S flight - considering only the images acquired along N-S direction, 3) E-W flight - considering only the images acquired along E-W direction, 4) sparse - considering one strip every two along both directions.

3.2.1 Pix4D Mapper Professional

Parrot and P4 have developed together the optimal workflow for Bebop UAS image processing. In particular, this software can

easily interpret the raw data acquired by the drone and has a good quality approximate calibration of the Sunny camera stored in its own database. The images have been processed with Pix4D Mapper Pro (version 4.0.25) maintaining the highest possible resolution for the initial processing (computation of the intrinsic and extrinsic parameters), thus allowing to perform the tie points search at a full-image resolution. It is worth noticing that P4 is capable of directing the tie-points search by creating a sky-mask. This task allows speeding up the tie-points search, while the matching is facilitating using the data provided by the camera geo-localization. Furthermore, P4 allows expressing the GCPs coordinates in a different reference system, with respect to the one used for camera geo-localization (e.g. local reference system for the GCPs and WGS84 for the telemetry). This is possible because the software computes the roto-translation between the difference reference systems, once the sparse point cloud has been generated and few GCPs have been measured.

The dense point clouds have been generated considering images with dimensions equal to half of those of the original images, using high point density, resulting in a points clouds composed by 17132k points. The correspondent 3D meshed models have been created with a medium resolution, while the orthophotos have been generated with a pixel size equal to 0.05 m.

3.2.2 Agisoft PhotoScan Professional

The acquired images have been also processed with Agisoft PhotoScan Professional (version 1.3.4), following the standard workflow. The dataset was used considering the full image resolution for the extrinsic parameters estimation (correspondent to the high alignment quality of PS). However, some problems have emerged because of the non-coherent value of pixel size computed from the EXIF file with respect to the nominal focal length (731 instead of 1403 pixel). For this reason, previously estimated Internal Orientation (IO) parameters have been used as approximated values. The Agisoft Lens application was discarded for camera calibration because of its inaccurate results. Finally, the computed high quality solution has been refined using the 'optimize stage' embedded in PS. In order to obtain results comparable to those obtained using P4, the rolling shutter correction has been performed too. According to Vautherin et al. (2016) the rolling shutter correction is performed directly on Bebop; however, for PS it was necessary to over-parametrize the Brown's calibration model enabling also the rolling shutter correction. PS is not capable of managing different reference systems between the GCPs and the camera geo-localization. For this reason, the local coordinates of the GCPs have been transformed in the WGS84 reference system, allowing to use the camera geo-localization data for improving and speeding the image selection during the tie-points detection and matching phases (i.e. use of PS reference mode). In fact, PS allows to project geographic coordinates onto cartographic ones (and vice versa), but it is not possible to process together data expressed in local coordinates (es. GCPs) and geo-localization information expressed in a global reference system.

The dense cloud was generated maintaining the 'high' quality, downgrading the images by a factor 4. This choice was made to guarantee the same level of resolution used in P4; however, we verified that this could be very demanding in term of the time requested for the computation, mainly because the last step of the dense point cloud generation is not optimized for GPU processing. The result is a point cloud composed by 14393k points. In addition, the PS orthoimage has been generated with a pixel size equal to 0.05 m.

3.2.3 Matlab 2017b + Agisoft PhotoScan Professional

The last tested approach consisted in a 2-step procedure, based on the creation of undistorted images using MAT and their

processing in PS. In Matlab 2017b, new functions that allows calibrating a fisheye camera and correcting its images from lens distortions have been embedded (see Section 2 for details about the mathematical model). Firstly, a series of images of the MAT calibration checkerboard have been acquired, by moving the Bebop in front of the panel itself and rotating the UAS by 90°, both clockwise and counter clockwise. Then the *estimateFisheyeParameters* function has been used to determine the camera intrinsic parameters, according to the model presented in Scaramuzza et al. (2006). They were used to correct all the images acquired during the survey, using the *undistortFisheyeImages* function. The output of this function are the undistorted images (with size equal to the one of the original images), as well as the parameters of a virtual perspective camera that produce such images. The IO parameters that describes such virtual camera have been refined by performing a self-calibration, in order to take into account the effect of the vibration that the camera suffers during take-off and landing. The generated images covers a smaller area, because during the resampling phase the higher distorted parts (located near the borders) are discarded. As one can notice (see Fig.2), the straight lines are mapped as straight lines because the undistorted images generated by MAT obey to the classical perspective projection model.

The image overlapping remains quite high (>9), thus guaranteeing anyway a complete 3D model. The dense cloud has been generated using the high resolution, resulting in a cloud composed by 16275k points. The orthoimage was generated with a pixel size equal to 0.05 m, as already done for all the others software packages.

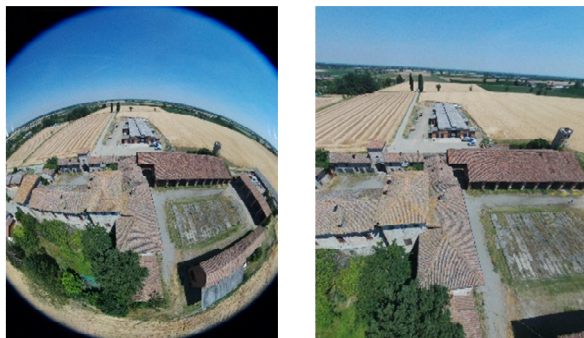


Figure 2. Image acquired with Bebop UAS (left) and the same image after correction in MAT (right)

3.3 Bundle block adjustment results

The quality of the photogrammetric solutions was evaluated in terms of residuals on the CPs, but also considering the value of the intrinsic parameters common to all the distortion models, namely the focal length and the position of the principal point. Concerning the distortion parameters, it was not possible to perform a direct comparison, because of the different models implemented within the different software packages. In Table 2 the residual of the CPs after bundle block adjustment are shown, for all the considered scenarios.

The obtained results are quite promising and in line with the expected accuracy. P4 ensures high quality results, but it suffers more when images acquired only along a specific direction are used. The 2-step MAT calibration approach ensures high accuracy in all the considered scenarios, with precision always below one GSD. PS seems to be a little more problematic because it requires a good approximation of the calibration parameters (mainly because the value of the pixel size and the one of the focal length read from the EXIF file are not coherent).

Software	Block configuration	CP residual [m]		
		X	Y	Z
P4	Grid	0.014	0.019	0.013
P4	N-S	0.032	0.054	0.010
P4	E-W	0.030	0.025	0.015
P4	Sparse	0.019	0.016	0.014
PS	Grid	0.024	0.017	0.037
PS	N-S	0.043	0.036	0.065
PS	E-W	0.018	0.015	0.020
PS	Sparse	0.020	0.026	0.052
MAT+PS	Grid	0.017	0.022	0.010
MAT+PS	N-S	0.014	0.033	0.013
MAT+PS	E-W	0.018	0.019	0.013
MAT+PS	Sparse	0.020	0.028	0.015

Table 2. CPs residuals after bundle block adjustment

The PS results are in line with the requested tolerance; however, they are slight worse with respect to other software packages, especially along the height direction. Probably, this result can be refined by improving the external camera calibration stage. However, because Agisoft Lens does not seem to work very well, it is necessary to rely on a calibration method that uses the same equidistant projection model used by PS. In other words, it is not possible to calibrate in one software (e.g. MAT) and use the results and the original images directly in a second software (e.g. PS), when the implemented projection and distortion models are different. In Table 3, the estimated intrinsic parameters are shown, considering only those that are common between the different models.

Software	Focal length [pixel]	Principal point x [pixel]	Principal point y [pixel]
P4	1423.33±0.06	14.74±0.02	-3.17±0.02
PS	1425.01±0.43	-2.47±1.50	-11.46±0.18
MAT	1416.13±1.27	8.67±1.04	-7.85±1.08

Table 3. Estimated intrinsic parameters. For P4 and PS they correspond to those of self-calibration, while with MAT an a-priori calibration was done using a checkerboard panel

It is quite evident that there is a certain level of variation among the different parameters, especially for the position of the principal point (along the x-axis it changes its sign in case of PS estimation). It is important to notice that the estimated parameters could be highly correlated with the distortion ones, and the variance and covariance matrix is not available for all the performed estimation.

3.4 Point Cloud Analysis

As a further analysis, the quality of both sparse and dense point cloud have been investigated. The analysis discussed below are presented only for the gridded flight configuration, because from the results presented in previous Section 3.3, it has emerged that they are not significantly different varying the image block configuration.

Concerning the sparse point clouds (i.e. the tie-points matching), they have been compared to verify the spatial distribution of the homologous points (e.g. possible areas where the matching is more difficult for one software because of the specific texture). To this regard, the data have been gridded (with a grid spacing equal to 50 GSD, corresponding to 2.5 m) and for each voxel the frequency of the contained tie-points has been evaluated (see Fig.3). It is worth noticing that the number of matched points is higher for P4 (339k points for P4, 85k points for PS and 112k

points for MAT+PS), thus resulting in a higher absolute frequency, even if P4 cloud is the only one that shows a hole in correspondence of the main entrance. For all the software packages, there are few false matching (e.g. located under the ground or in the sky).

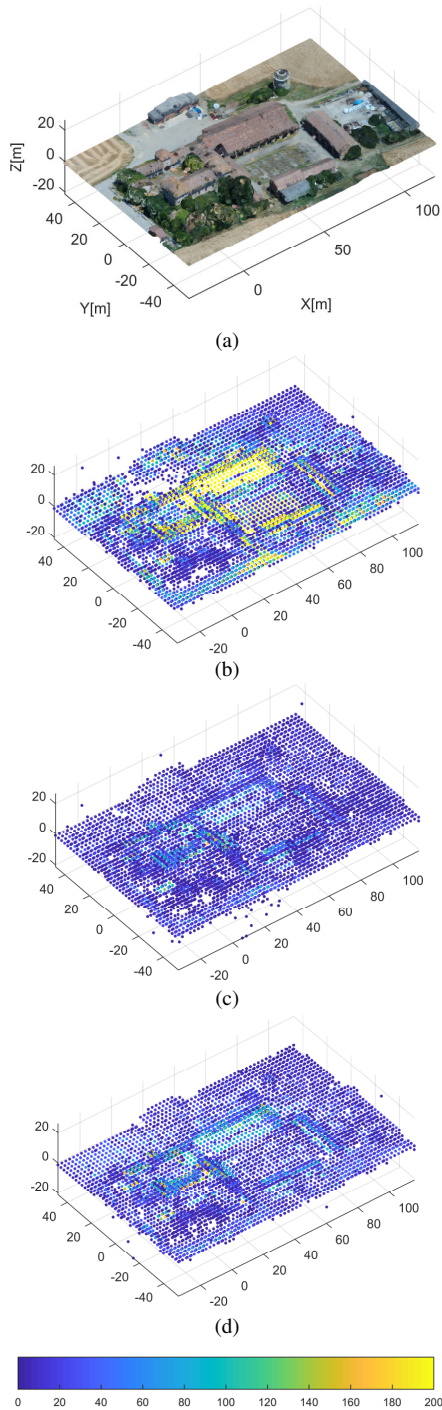


Figure 3. a) 3D model of the surveyed area
 Frequency of the matched points in voxel with a size equal to 50 GSD: b) P4 - c) PS - d) MAT+PS. The scale colored bar represents the absolute frequency of the matched points

Considering the dense clouds, the data generated with the three different proposed approaches have been compared with a reference laser point cloud, acquired by means of a Leica MS60 MultiStation. The reference cloud has been acquired by

stationing on two vertexes of the geodetic network (respectively point 100 and point 200, see Fig.1), using a horizontal angular resolution of $0^{\circ}.1432$ and a vertical resolution equal to $0^{\circ}.0573$, thus resulting in a spatial resolution equal to 0.05 and 0.02 m, at an object distance equal to 20 m. The reference point cloud is composed by $3.3 \cdot 10^6$ points, with a theoretical accuracy of 1 mm at 50 m. The point clouds generated using PS (namely PS and MAT+PS scenario) have been manually edited to remove points located under the ground, obviously due to false matching. All the dense clouds have been edited to remove vegetated areas. The differences among the analyzed datasets have been computed using the M3C2 plugin (Lague, 2013), implemented in the Cloud Compare Software (v 2.7.0). All the points with differences (in absolute terms) greater than 0.20 m have been considered outliers, and removed from further analysis. The mean, the standard deviation and the RMS of the differences obtained from the comparison of each dataset with the reference scan are shown in Table 4. The results are divided between horizontal (H1 and H2) and vertical planes (V1 and V2), see Fig.4.

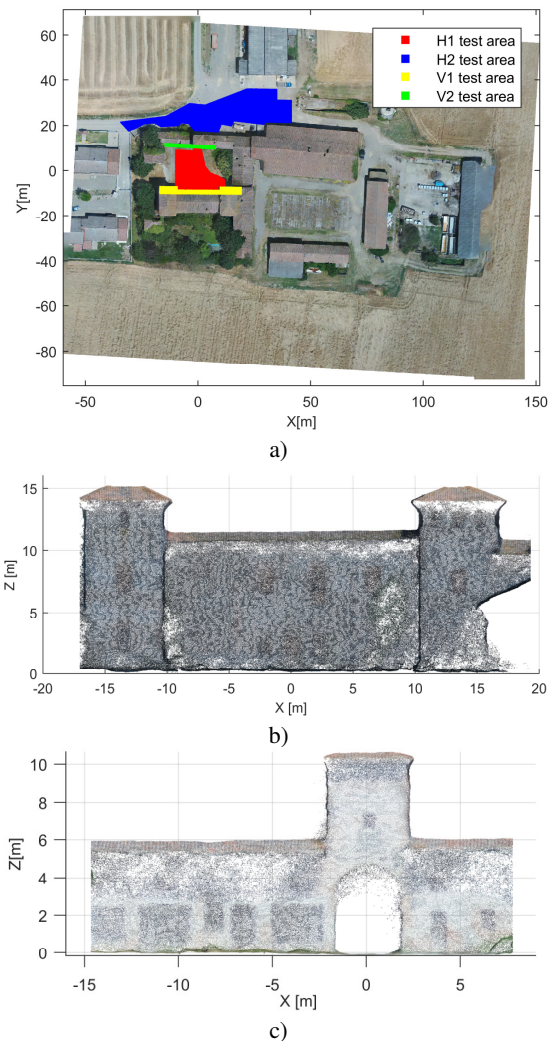


Figure 4. a) The four areas considered for comparing generated points clouds with the reference MS scan b) front considered for V1 c) front considered for V2

It is worth noticing, that the computed differences are not significant for both the vertical areas (namely V1 and V2), instead they are of the order of the GSD for both the horizontal areas (H1 and H2). This could be partially explained considering the different accuracy of the altimetric survey, with respect to the

planimetric one. The computed standard deviations are higher for the vertical surfaces compared to the horizontal ones, as well as the number of outliers. This is probably due to higher level of roughness that characterizes V1 and V2, because of the presence of doors, windows, building corners, vegetation etc. The Root Mean Squares Errors (RMSEs) are of the order of the GSD for all the analyzed cases, apart from PS for V1 and V2 datasets (about 1.5 GSD) and PS for H1 dataset (0.8 GSD).

Area (# points)	Software	Outliers %	Mean [m]	St. Dev [m]	RMSE [m]
H1 (1350k)	P4	0.5	-0.057	0.019	0.060
	PS	0.6	-0.036	0.015	0.039
	MAT+PS	0.7	-0.051	0.013	0.052
H2 (442k)	P4	0.1	-0.056	0.011	0.057
	PS	0.1	-0.049	0.031	0.058
	MAT+PS	0.1	-0.056	0.011	0.057
V1 (776k)	P4	6.5	0.000	0.054	0.054
	PS	16.8	0.007	0.085	0.086
	MAT+PS	11.8	0.004	0.065	0.065
V2 (293k)	P4	11.7	-0.008	0.050	0.050
	PS	15.7	0.004	0.072	0.072
	MAT+PS	14.0	-0.006	0.057	0.057

Table 4. Statistics of the differences between the dense clouds and the reference laser scan, for the four different analysed areas

4. CONCLUSIONS

In this paper, an analysis of the use of fisheye oblique images (acquired with a Parrot Bebop 2 UAS) for 3D modelling using photogrammetric commercial software is presented.

Different fisheye mathematical models have been exploited, paying specific attention to those implemented in Pix4D Mapper Pro, Agisoft PhotoScan and to the new fisheye calibration and image correction functions embedded in Matlab 2017b. Then, an experimental test on the use of the Bebop 2 UAS has been presented, analyzing points of strength and weaknesses of such system, together with the achievable metric accuracy of the photogrammetric solution and of the derivable products (e.g. sparse and dense clouds).

From the analysis conducted on the mathematical models has emerged that different formulations are implemented in different software packages, however there is no particular hint to define which one is the most valid. They are based on different assumption and projection geometry, resulting that they are not interoperable one another. This means that it is not possible to perform the camera calibration using one software and used the estimated parameters in a second one, as can be easily done with perspective/pinhole cameras.

For all the considered software packages, the obtained results are quite satisfactory considering the great differences between rectilinear images and the fisheye ones, acquired by the Bebop UAS. The residuals of the checkpoints are below the GSD (equal to 0.05 m considering the direction normal to the sensor axis for a flight height equal to 35 m) in all the directions, for all the tested flight configurations. These results have been confirmed using the two evaluated approaches, namely using the original fisheye images in commercial software packages (P4 and PS) or previously calibrating the camera in MAT and using the output-undistorted images in PS.

The use of oblique UAS fisheye images has clearly emerged as an interesting solution for creating a complete 3D model of complex structure of buildings, courtyards etc., and reconstructing vertical surfaces too. In fact, the use of UAS with

an oblique camera allows seeing also the building façades, which typically cannot be reconstructed with standard nadiral flights. The main drawbacks of Bebop UAS is the small dimension of the sensor, which results in a quite wide GSD. If higher quality results are request, it is necessary to fly to lower altitude, increasing consequently the number of images to be processed. Another drawback of the used UAS is that the images are characterized by a large field of view (180°), which in principle can reduce the number of images to be acquired. However, they are characterized by a high variability of the GSD, meaning that it is important to ensure a high overlapping in order to compensate for this effect, without affecting the quality of the final 3D model.

Different flight configurations have been analyzed with the aim to evaluate their impact on the final photogrammetric block accuracy. In all the considered scenarios, the residuals on the checkpoints were of the order of the GSD, in all directions, reading that the impact of the flight configuration could be neglected in case of very high overlapping. This hypothesis is confirmed by the fact that only slight increment of the residuals can be underlined for N-S scenario (along this direction the strips were shorter). However, it is worth noticing that the grid flight configuration results in a largest and more detailed dense cloud, which allows to better reconstruct all the building façades in case of complex structures, such as the investigated ones. Comparing the size of the sparse point clouds (i.e. the number of matched object points used for the bundle block adjustment) generated by the different software packages, it clearly emerges that P4 creates a bigger dataset, even if its spatial distribution is more variable and few holes are present. The point clouds generated by PS are sparser, also when the used images are those previously undistorted in MAT. These differences in the tie-points distribution and density are due to the different algorithms used for homologous points search and matching. However, the evaluated software packages are commercial ones, so it is not possible to know which is the interest operator effectively used and how it is implemented. The generated dense clouds have been compared to a reference laser scan acquired by means of a Leica MS60 MultiStation, showing differences no larger than 0.05 m. Finally, the orthoimages have been compared along horizontal planes, even if this analysis has been not here discussed because of the limited available space. The results are quite good, considering the high level of distortion that affected the Bebop images and that the orthoimage pixel size is equal to the GSD normal to the optic axis (more or less double than the nadiral GSD), however few problems have been noticed in correspondence to the edges, especially when P4 is used.

ACKNOWLEDGMENTS

The authors would like to thank Ing. Daniele Passoni for the realization of the UAS flights, and Mario Bulgarini, who discussed is Bachelor thesis on this topic, graduating in Civil Engineering at Politecnico di Milano.

REFERENCES

Accardo, D., Fasano, G., Forlenza, L., Moccia, A., Rispoli, A., 2013. Flight test of a radar-based tracking system for UAS sense and avoid. In: *IEEE Transactions on Aerospace and Electronic Systems*, 49(2), pp. 1139-1160. doi.org/10.1109/TAES.2013.6494404

- Barazzetti, L., Previtali, M., Roncoroni, 2017. F., Fisheye lenses for 3D modeling: evaluations and considerations, In: *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XLII-2/W3, pp. 79-84. <https://doi.org/10.5194/isprs-archives-XLII-2-W3-79-2017>.
- Bolognesi, M., Furini, A., Russo, V., Pellegrinelli, A., Russo, P., 2015. Testing the low-cost RPAS potential in 3D cultural heritage reconstruction. In: *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XL-5/W4, pp. 229-235. <https://doi.org/10.5194/isprsarchives-XL-5-W4-229-2015>.
- Brown, C., 1971. Close-range camera calibration. *Photogrammetry Engineering and Remote Sensing*, 37(8), pp. 855-866.
- Courbon, J., Mezouar, Y., Eckt, L., Martinet, P., 2007. A generic fisheye camera model for robotic applications. In: *Intelligent Robots and Systems. IROS 2007. IEEE/RSJ International Conference on*, pp. 1683-1688. doi.org/10.1109/IROS.2007.4399233.
- Covas, J., Ferreira, V., Mateus, L., 2015. 3D reconstruction with fisheye images strategies to survey complex heritage buildings. In: *Digital Heritage*, Vol. 1, IEEE, pp. 123-126. doi.org/10.1109/DigitalHeritage.2015.7413850.
- de Miguel Molina, B., Oña, M. S., 2018. The Drone Sector in Europe. *Ethics and Civil Drones*, Springer, Cham, pp. 7-33.
- Eisenbeiss, H., 2004. A mini unmanned aerial vehicle (UAV): system overview and image acquisition. In: *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XXXVI-5/W1, pp. 1-7.
- El-Hakim, S. F., 1986. Real-time image metrology with CCD cameras. *Photogrammetry Engineering and Remote Sensing*, 52.11, pp. 1757-1766.
- Fagandini, R., Federici, B., Ferrando, I., Gagliolo, S., Pagliari, D., Passoni D., Pinto, L., Rossi, L., Sguerso, D., 2017. Evaluation of the Laser Response of Leica Nova Multistation MS60 for 3D Modelling and Structural Monitoring. In: *International Conference on Computational Science and Its Applications*, pp. 93-104. [doi.org/ 10.1007/978-3-319-62401-3_8](https://doi.org/10.1007/978-3-319-62401-3_8).
- Gagliolo, S., Fagandini, R., Federici, B., Ferrando, I., Passoni, D., Pagliari, D., Pinto, L., Sguerso, D., 2017. Use of UAS for the conservation of historical buildings in case of emergencies. In: *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XLII-5/W1, pp. 81-88. <https://doi.org/10.5194/isprs-archives-XLII-5-W1-81-2017>.
- Huang, Y. P., Sithole, L., Lee, T. T., 2017. Structure From Motion Technique for Scene Detection Using Autonomous Drone Navigation. In: *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, pp. 1-12.
- Hughes, C., McFeely, R., Denny, P., Glavin, M., Jones, E., 2010. Equidistant fish-eye perspective with application in distortion centre estimation. *Image and Vision Computing*, 28(3), pp. 538-551.
- Kannala, J., Brandt, S. S., 2006. A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses. In: *IEEE transactions on pattern analysis and machine intelligence*, 28(8), pp. 1335-1340. <http://doi:10.1109/TPAMI.2006.153>.
- Lague, D., Brodu, N., Leroux, J., 2013. Accurate 3D comparison of complex topography with terrestrial laser scanner: Application to the Rangitikei canyon (NZ). *ISPRS Journal of Photogrammetry and Remote Sensing*, 82, pp. 10-26.
- Perfetti, L., Polari, C., Fassi, F., 2017. Fisheye photogrammetry: tests and methodologies for the survey of narrow spaces. In: *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XLII-2/W3, pp. 573-580. <https://doi.org/10.5194/isprs-archives-XLII-2-W3-573-2017>.
- Piras, M., Di Pietra, V., Visintini, D., 2017. 3D modeling of industrial heritage building using COTSs system: Test, limits and performances. In: *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XLII, pp. 281-288.
- Remondino, F., Barazzetti, L., Nex, F., Scaioni, M., Sarazzi, D., 2011. UAV photogrammetry for mapping and 3d modeling—current status and future perspectives. In: *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XXXVIII-1/C22.
- Scaramuzza, D., Martinelli, A., Siegwart, R., 2006. A toolbox for easily calibrating omnidirectional cameras. In: *Intelligent Robots and Systems, IEEE/RSJ International Conference on*, pp. 5695-5701. doi.org/10.1109/IROS.2006.282372.
- Schneider, D., Schwalbe, E., Maas, H. G., 2009. Validation of geometric models for fisheye lenses. *ISPRS Journal of Photogrammetry and Remote Sensing* 64.3, pp. 259-266.
- Schwalbe, E., 2005. Geometric modelling and calibration of fisheye lens camera systems. In: *International Archives of Photogrammetry and Remote Sensing*. Vol. XXXVI-5, pp.1-8.
- Steffen, A., Förstner, W., 2005. Fish-eye-stereo calibration and epipolar rectification. *ISPRS Journal of photogrammetry and remote sensing*, 59(5), pp. 278-288.
- Taufik, A., Okamoto, S., Lee, J. H., 2016. 3D Mapping of an Environment around a River Using Image Processing of Photo Images Captured by a Multi-Rotor Drone and 3D Map Modeling Using a 3D Printer. In: *Proceeding ECBA-2016*, Vol. 127-Issue.5, pp. 1-9.
- Vautherin, J., Rutishauser, S., Schneider-Zapp, K., Choi, H. F., Chovancova, V., Glass, A., Strecha, C., 2016. Photogrammetric Accuracy and Modeling of Rolling Shutter Cameras. In: *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. III-3, pp. 1-8.
- Yan, J., Grasso, N., Zlatanova, S., Braggaa, R. C., Marx, D. B., 2017. Challenges in flying quadrotor unmanned aerial vehicle for 3D indoor reconstruction. In: *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XLII, pp 423-430. doi.org/10.5194/isprs-archives-XLII-2-W7-423-2017.