A Structure-from-Motion Approach Using UAV-based Imagery for Precision Agriculture Applications

F. He [a, *], W. Xiong [a], A. Habib [a]

[a] Digital Photogrammetry Research Group
Lyles School of Civil Engineering
Purdue University, 550 Stadium Mall Dr., West Lafayette, IN 47907, USA
– (he270, xiong43, ahabib)@purdue.edu

KEY WORDS: Structure from Motion, UAVs mapping, Bundle Adjustment, Precision Agriculture

ABSTRACT:

Nowadays, thanks to recent advances in sensor calibration and automated triangulation, accurate mapping using low-cost Unmanned Airborne Vehicles (UAVs) with RGB frame camera is gaining significant attention from researchers in the mapping and plant science fields. However, due to the fact that the crop fields usually have poor and/or repetitive textures, 3D image-based reconstruction using UAV images acquired for precision agriculture applications remains to be a challenging task. In this paper, a three-step Structure from Motion (SfM)-based procedure is proposed for accurate 3D reconstruction using UAV-based images. In the first step of the proposed procedure, two approaches, which take advantage of prior information regarding the flight trajectory, are proposed for reliable relative orientation in the presence of a high percentage of outliers. Then, image orientation recovery and sparse point cloud generation are simultaneously achieved through an image augmentation process. Finally, a global bundle adjustment involving Ground Control Points (GCPs) and check points is carried out to refine the estimated image EOPs relative to the mapping reference frame. In order to evaluate the feasibility of the proposed procedure, two real datasets, which are acquired with different configurations over a test agriculture field, are utilized. In addition, a comparative performance analysis between the proposed procedure and the Pix4D software is performed. The derived experimental results demonstrate that the proposed procedure is capable of providing accurate UAV-based 3D reconstruction for precision agriculture applications. It also presents superior performance when compared to the Pix4D software.

1. INTRODUCTION

In conventional photogrammetric mapping, large-area 3D reconstruction has been established using manned-airborne data acquisition platforms. Such image-based 3D reconstruction requires the knowledge of the Interior Orientation Parameters (IOPs) of the utilized camera, Exterior Orientation Parameters (EOPs) of the involved images, and corresponding points/features in the set of overlapping images. The IOPs can be derived from a camera calibration process. The EOPs of the involved imagery can be either derived through an indirect geo-referencing procedure using tie and control points or a direct geo-referencing process through the implementation of a GNSS/INS unit on-board the mapping platform. While the latter approach has the practical convenience of simplifying the geo-referencing process, it requires significant initial investment for the acquisition of a high-end GNSS/INS Position and Orientation System (POS) – especially, when seeking high level of reconstruction accuracy.

Instead of using conventional manned-photogrammetric data acquisition platforms, low-cost UAVs have recently emerged as a promising geospatial data acquisition system to satisfy the needs of wide range of applications (Tao and Li, 2007). Compared to manned-airborne systems, UAVs are more advantageous for small area mapping since they can be quickly deployed and flown at lower elevation resulting in more cost-effective 3D mapping at higher resolution. In the existing body of literature, SfM approaches and dense image matching techniques have been investigated. SfM automates the process of EOP recovery. Dense image matching, on the other hand, provides feature correspondences among overlapping images at the pixel level. In addition, thanks to recent advances in both low-cost direct geo-referencing systems as well as imaging sensors operating at different portions of the electromagnetic spectrum, utilizing UAVs, which are equipped with directly geo-referenced RGB-frame cameras, for precision agriculture has become an important application that is gaining significant attention from researchers in both mapping and plant science fields (Habib et al., 2016). It is worth mentioning that, although current commercial software (e.g., Pix4D and PhotoScan) can automate the process of image-based 3D reconstruction, utilization of such software for precision agriculture applications remains to be a challenging task. This is mainly due to the fact that agriculture fields usually contain poor and/or repetitive textures, which can severely impact the relative orientation recovery for UAV-based images. Therefore, this paper aims at proposing a transparent SfM-based system, which can be augmented with different user-defined constraints, for precision agriculture applications. To be more specific, this research will be focusing on the following issues:

1. Automated relative orientation recovery of UAV-based images in the presence of prior information regarding the flight trajectory,
2. Sparse image-based point cloud generation from multiple images acquired from low-cost Large-field-of-view (LFOV) cameras onboard UAV platforms,
3. Accuracy analysis of the derived 3D reconstruction through check point analysis, and
4. Comparison of the proposed approach against available commercial software, such as Pix4D.

* Corresponding author.

To address these issues, the utilized methodology is introduced in the next section. Afterwards, experimental results with real datasets are discussed. Finally, drawn conclusions as well as recommendations for future work are presented.

2. METHODOLOGY

In this section, the proposed SfM procedure for UAV-based 3D reconstruction while using a low-cost LFOV camera is introduced. Figure 1 illustrates the workflow of the proposed procedure. This procedure includes three steps. In the first step, the relative orientation parameters relating stereo-images are directly derived from conjugate point features. In the second step, the initial recovery of image EOPs is achieved. Specifically, a local coordinate frame is initially established, and then the EOPs of the remaining images are sequentially recovered in an incremental image augmentation process. Finally, in the third step, a global bundle adjustment, which integrates both Ground Control Points (GCPs) and Check Points, is carried out for indirect geo-referencing and accuracy analysis.

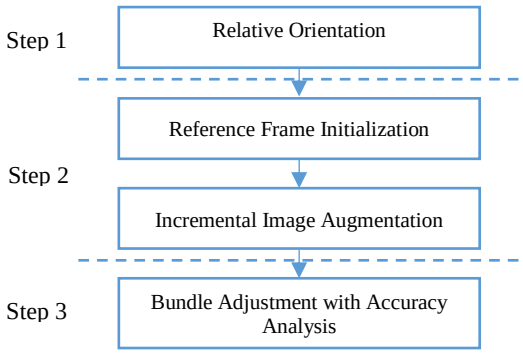


Figure 1. The proposed workflow for the SfM-based 3D reconstruction

2.1 Relative Orientation

In the past few decades, recovery of the Relative Orientation Parameters (ROPs) relating overlapping images has been investigated within both the photogrammetric and computer vision research communities (Habib and Kelley, 2001; Heipke, 1997; Zhang et al., 2011). In general scenarios, the IOPs of the utilized camera are usually assumed to be known. For a given stereo-pair, ROP estimation involves the derivation of five parameters, which include three rotation angles and two translation parameters (i.e., an arbitrary scale is assumed for the ROP estimation procedure). The most well-known approach for ROP recovery is based on the co-planarity constraint (Mikhail et al., 2001), where a least-squares adjustment is solved while using a minimum of five conjugate points. However, due to the nonlinear nature of the co-planarity model, approximate values for the unknowns, which are refined through an iterative process, have to be available. Establishing good-quality approximations can be challenging for situations where the mapping platform exhibits excessive maneuvers between the data acquisition epochs (e.g., close range mapping applications). To address such a challenge, several closed-form solutions, which do not require approximations, for ROP recovery have been developed, such as the eight-point and the five-point algorithms (Hartley, 1997; Longuet-Higgins, 1987; Nistér, 2004). Although such closed-form solutions have been widely used in various applications, mitigating the impact of outliers in automatically matched points remains to be a challenging task. In order to deal with UA-based images acquired for precision agriculture applications, two approaches are proposed to derive reliable estimation of ROPs in the presence of high percentage of matching outliers. The proposed approaches have following characteristics:

1. Both proposed approaches take advantage of prior information regarding the flight trajectory, which can be derived from the designed mission plan and/or geo-referencing information from an onboard GNSS/INS unit.

2. The first approach is based on a two-point closed-form solution. It assumes that the UAV is moving at constant flying height while operating a nadir-looking (i.e., we are dealing with vertical images that are captured from the same flying height). Starting from such assumptions, the nine elements of the Essential Matrix, which contains all information regarding the ROPs relating two overlapping images, can be reduced to four with an additional constraint among them. Meanwhile, considering the fact that the relative orientation can be only established up to an arbitrary scale, the four elements of the Essential Matrix can be derived using a minimum of two point correspondences. In addition, this approach can be incorporated within a RANSAC framework for outlier removal among the initial conjugate point pairs.

3. The second approach is an iterative five-point approach. It starts with a linearization process of the co-planarity constraint starting from prior information regarding the ROPs relating the stereo-pair in question to derive a set of linear equations in the unknown corrections to the approximate ROP values. A minimum of five conjugate pairs are required to derive an estimate of the ROPs through an iterative procedure until a convergence criterion is achieved. In this approach, a built-in outlier removal process is adopted within the iterative procedure by imposing constraints on the normalized image coordinates according to epipolar geometry.

4. In practice, due to wind conditions and the light-weight of UAVs, the acquired UAV images are expected to show significant deviations from the assumption of the proposed two-point approach (i.e., the images are acquired with the camera's optical axis pointing in the vertical direction and at the same flying height). Therefore, a hybrid approach that starts with the two-point approach and uses the estimated initial values for the iterative five-point approach can be adopted when dealing with stereo-images that exhibit significant variations from the designed flight plan.

2.2 Initial Recovery of Image EOPs

Once the ROPs of all possible stereo-pairs are estimated, an incremental approach, which is developed by He and Habib (2014), is adopted for the initial recovery of the image EOPs. This incremental approach is initiated by defining a local coordinate frame. Then, all the images are sequentially augmented into a final image block.

The local coordinate system is established using an image triplet that satisfies both a large number of corresponding points and good compatibility configuration. The first condition can be easily satisfied by maximizing the total number of corresponding points within the image triplet. The second condition is evaluated through a method referred to as compatibility analysis, which was developed by He and Habib (2014).

Once the local coordinate frame is established, the remaining images are sequentially augmented into the final image block or

trajectory. The proposed approach utilizes a geometric structure, which is referred to as the reverse tree structure for the augmented images (Martinec and Pajdla, 2007), as the basis for the EOP recovery. In this research, the proposed reverse tree-structure-based approach incorporates a linear approach for the estimation of the image EOPs (He and Habib, 2014). In order to reduce the effects of error propagation, at each step of the image incremental augmentation, only the image that exhibits the highest compatibility with the previously referenced imagery is selected and referenced into the pre-defined local coordinate system.

2.3 Global Bundle Adjustment

Since the derived sparse point cloud from the SfM approach is only defined in an arbitrary local coordinate system, an absolute orientation process has to be applied for defining the derived point cloud as well as the estimated image EOPs relative to the mapping reference frame. In this research, correspondences between the GCPs and the derived sparse point cloud are first manually identified. Then, a 3D similarity transformation process is performed to evaluate approximate parameters (i.e., scale factor, three translation parameters, and three rotation angles) relating the two involved coordinate systems.

Now that the 3D similarity transformation has been applied, the accuracy analysis of the derived UAV-based 3D reconstruction results can be conducted through a global bundle adjustment integrating the GCPs and check points. In the implemented bundle adjustment process, GCPs are established for absolute orientation/datumn definition, and check points are established for accuracy evaluation. In this research, besides the derived conjugate points from the SfM approach, tie points corresponding GCPs and check points are manually identified. Then, these image measurements are integrated into the bundle adjustment for a global refinement. Finally, the residuals at each check point are computed. Figure 2 illustrates the inputs and outputs of the global bundle adjustment process.

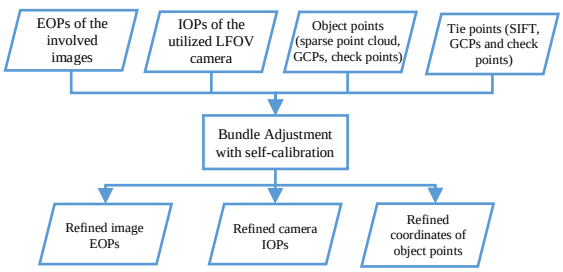


Figure 2. The proposed global bundle adjustment process for accuracy analysis

3. EXPERIMENTAL RESULTS

The main objective of the experimental results is illustrating the feasibility of the proposed SfM-based procedure, which aims at 3D reconstruction using UAV-based images for precision agriculture applications. Therefore, we conduct several experimental tests using real image datasets.

3.1 Dataset Description

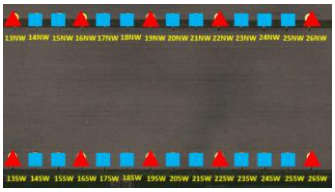
The test site involved in the experiments covers an agriculture field at Purdue Agronomy Center for Research and Education (ACRE) in Wester Lafayette, United States. Two real datasets are acquired by a multi-rotor DJI Phantom2 UAV with GoPro Hero 3+ Camera over the test site. The internal characteristics of the GoPro camera has been estimated through a calibration procedure similar to the one proposed by He and Habib, (2015). One should note that, for the utilized UAV platform, the GoPro camera is mounted on a gimbal to ensure that images are acquired with the camera's optical axis pointing in the nadir direction. The main characteristics of the two real datasets are described as below.

Dataset 1 is comprised of 569 images that are captured from a flying height of approximately 15 meters with the utilized multi-rotor UAV moving at a speed of roughly 8 m/s (See Figure 3(a) for sample image). The overlap and side lap percentages for the acquired images are approximately 60%.
Dataset 2 includes 1,191 images acquired with a 5 m/s speed of the same UAV platform at a flying height of almost 15 meters (See Figure 3(b) for sample image). The overlap and side lap ratio among the acquired images are almost 80%.

(a) (b)

Figure 3. (a) A sample image from real dataset 1, and (b) a sample image from real dataset 2

In order to evaluate the derived 3D reconstruction results from the two real datasets, 10 GCPs and 18 check points are established. Both GCPs and check points are surveyed by an RTK GPS with an approximate accuracy of $\pm 2\ cm$. The configuration of the utilized GCPs and check points are shown in Figure 4.



(a)

(b)

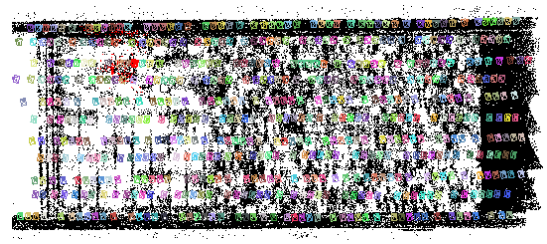
Figure 4. (a) Configuration of the GCPs and check points, and (b) a sample image of the targets established for both GCPs and check points in the test field

3.2 Accuracy Analysis

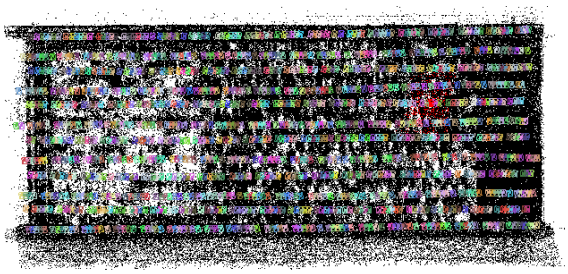
For each real dataset, by applying the proposed SfM-based procedure, the EOPs of involved images are recovered, and a sparse point cloud is simultaneously generated. Figure 5 illustrates the generated sparse point cloud as well as the positions and orientation of all involved images for the two real datasets. Afterwards, a 3D similarity transformation process is

applied using at least three GCPs with good spatial distribution for deriving initial values of EOP and ground coordinates relative to the reference frame. Finally, the global bundle adjustment with GCPs and check points is carried out to refine the derived image EOPs, object coordinates, and camera IOPs. One should note that, since the number of derived tie points can vary a lot from different approaches (e.g., the proposed SfM-based procedure usually provides hundreds of thousands of tie points, while only hundreds of tie points can be identified for the GCPs and check points), distinct variances are considered for different types of tie points in the global bundle adjustment. More specifically, in this research, the standard deviation of SfM-based tie points is set to 7 pixel; the standard deviation of tie points corresponding to the GCPs and check points are set to 0.5 pixel.

Once the global bundle adjustment process is completed, the derived square root of a posteriori variance (σ_0) can be computed. In this research, the derived σ_0 values from the two real datasets (0.00137 mm and 0.00128 mm, respectively) are all smaller than one pixel size (0.00155 mm), which indicates a good precision of the derived UAV-based 3D reconstruction. In the meantime, in order evaluate the accuracy of the derived 3D reconstruction results in the mapping coordinate system, the root-mean-square error (RMSE) for the check points is investigated. The derived RMSE values are presented in Table 1, where one can observe that, for both real datasets, the RMSE values for X and Y coordinates are 0.01 m, while the RMSE value for Z coordinate is 0.04 m. Such results indicate that the UAV-based 3D reconstruction achieves a good accuracy in the object space.



(a)



(b)

Figure 5. Reconstructed sparse point cloud and positions and orientation of all involved experimental images from (a) real dataset 1 and (b) real dataset 2.

	Real Dataset 1	Real Dataset 2
RMSE_X (m)	0.01	0.01
RMSE_Y (m)	0.01	0.01
RMSE_Z (m)	0.04	0.04
Total_RMSE (m)	0.02	0.02

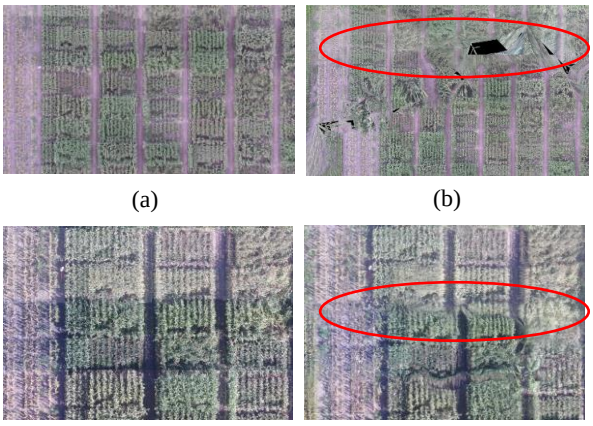
Table 1. The derived RMSE values at check points from real datasets 1 and 2

3.3 Comparison with Pix4D

The objective of the last stage of the experimental results is a comparative performance analysis between the proposed SfM-based procedure and the Pix4D software. Therefore, qualitative and quantitative evaluation methodologies need to be established. In this research, the quantitative analysis is performed by comparing the number of images, whose EOPs have been successfully recovered from the proposed procedure and the Pix4D software. Table 2 reports the number of estimated image EOPs from the real datasets. One can observe that the proposed SfM-based procedure provides EOP estimation for all images (569, and 1,191 images, respectively) from the two real datasets, while only 487 out of 569 and 1100 out of 1191 image EOPs are recovered from Pix4D. Such result demonstrates that the proposed SfM-based procedure is capable of providing more reliable EOP estimation when dealing with UAV images acquired over an agricultural field. In the meantime, the qualitative evaluation is conducted by evaluating the RGB-based orthophoto mosaic of the entire test field using the two real datasets. In this research, a DEM is initially interpolated from the sparse point cloud that is derived from either the proposed SfM procedure or Pix4D. Then, the DEM together with the bundle adjustment-based EOPs and the GoPro IOPs are used to generate the RGB-based orthophoto. Figure 6 illustrates the cropped RGB-based orthophotos that are respectively generated from the proposed SfM-based procedure and Pix4D. Through closer visual inspection on the derived orthophotos, one can observe obvious gaps and discrepancies in the RGB-based orthophoto generated by the Pix4D software (See Figures 6b and 6d). The orthophoto generated from the proposed procedure, on the other hand, shows good quality (See Figures 6a and 6b). Such results also indicate that, compared to Pix4D, the proposed procedure has superior performance when dealing with sets of agricultural field images acquired by the utilized UAV platform.

	Total Number of Images	Number of Images (The Proposed Procedure)	Number of Images (Pix4D)
Real Dataset 1	569	569	487
Real Dataset 2	1,191	1,191	1,100

Table 2. Number of recovered image EOPs for datasets 1 and 2



(a) (b)

(c) (d)

Figure 6. (a) and (b) are cropped from RGB-orthophotos generated from real dataset 1 using the proposed procedure and Pix4D; (c) and (d) are cropped from RGB-orthophotos generated from real dataset 2 using the proposed procedure and Pix4D.

4. CONCLUSIONS AND RECOMMENDATIONS FOR FUTURE WORK

In this paper, we proposed a fully automated procedure for UAV-based 3D reconstruction, which can be used for precision agriculture applications. Then, accuracy analysis is performed on the derived point clouds. Besides, a comparative performance analysis between the proposed procedure and Pix4D is carried out. For the experimental tests, two real datasets acquired by a DJI Phantom 2 UAV with a GoPro Hero 3+ camera are utilize. The derived experimental results from accuracy analysis using GCPs and check points indicate that the 3D reconstruction can be achieved with a good level accuracy. In addition, by comparing the derived results from the proposed procedure and Pix4D, we can draw the conclusion that the proposed procedure has better performance when dealing with UAV images acquired for precision agriculture applications.

It is important to note that, in this research, we only tested our proposed procedure on one low-cost UAV platform. For future work, more tests will be carried out on images acquired from different UAV platforms with different flight configurations. Besides, automated identification of GCPs and check points will be investigated as well.

ACKNOWLEDGEMENT

The information, data, or work presented herein was funded in part by the Advanced Research Projects Agency-Energy (ARPA-E), U.S. Department of Energy, under Award Number DE-AR0000593. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

REFERENCES

Habib, A., Kelley, D., 2001. Automatic relative orientation of large scale imagery over urban areas using modified iterated Hough transform. ISPRS J. Photogramm. Remote Sens. 56, 29–41.

Habib, A., Xiong, W., He, F., Yang, H.L., Crawford, M., 2016. Improving Orthorectification of UAV-Based Push-Broom Scanner Imagery Using Derived Orthophotos From Frame Cameras. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 1–15.

Hartley, R.I., 1997. In defense of the eight-point algorithm. Pattern Anal. Mach. Intell. IEEE Trans. On 19, 580–593.

He, F., Habib, A., 2015. Target-based and Feature-based Calibration of Low-cost Digital Cameras with Large Field-of-view, in: ASPRS 2015 Annual Conference, Tampa, Florida.

He, F., Habib, A., 2014. Linear approach for initial recovery of the exterior orientation parameters of randomly captured images by low-cost mobile mapping systems. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. 40, 149.

Heipke, C., 1997. Automation of interior, relative, and absolute orientation. ISPRS J. Photogramm. Remote Sens. 52, 1–19.

Longuet-Higgins, H.C., 1987. A computer algorithm for reconstructing a scene from two projections. Read. Comput. Vis. Issues Probl. Princ. Paradig. MA Fisch. O Firschein Eds 61–62.

Martinec, D., Pajdla, T., 2007. Robust rotation and translation estimation in multiview reconstruction, in: Computer Vision and Pattern Recognition, 2007. CVPR'07. IEEE Conference on. IEEE, pp. 1–8.

Mikhail, E.M., Bethel, J.S., McGlone, J.C., 2001. Introduction to modern photogrammetry. John Wiley & Sons Inc.

Nistér, D., 2004. An efficient solution to the five-point relative pose problem. Pattern Anal. Mach. Intell. IEEE Trans. On 26, 756–770.

Tao, C.V., Li, J., 2007. Advances in mobile mapping technology. CRC Press.

Zhang, Y., Huang, X., Hu, X., Wan, F., Lin, L., 2011. Direct relative orientation with four independent constraints. ISPRS J. Photogramm. Remote Sens. 66, 809–817.