# INTEGRATING UAV-BORNE AND AERIAL-BORNE IMAGERY FOR GENERATION OF 3D BUILDING MODELS

Min-Lung Cheng[1], Toshiaki Satoh[1, 2] and Masashi Matsuoka[1]
[1]Tokyo Institute of Technology, 4259-G3-2 Nagatsuta, Midori-ku, Yokohama, Kanagawa, 226-8502 Japan,
Email: cheng.m.ab@m.titech.ac.jp
Email: matsuoka.m.ab@m.titech.ac.jp
[2]Pasco Corporation, 1 Chome-1-2 Higashiyama, Meguro, Tokyo, 153-0043 Japan,
Email: tuoost7017@pasco.co.jp

**KEY WORDS:** Spatial Data, 3D building models, Unmanned Aerial Vehicle, Structure-from Motion, Photogrammetry

**ABSTRACT:** 3D digital city is one of the emphasized topics in geomatics and computer vision for its various applications. The functionalities of virtual systems can be applied to simulate a future city or post-disaster monitoring and management by spatial data processing, analysis and visualization. Nowadays, a lot of platforms such as satellite, aircraft and unmanned aerial vehicle (UAV) are widely performed to collect spatial data. Comparing to high-end digital photogrammetry systems, UAV is a low-cost platform to acquire spatial information in the real world. Conventionally in the spatial data processing, UAV-borne photos and high-altitude aerial images are processed separately to their requirements. Thus for generation of 3D building models, there may be distinct geometrics representing the same object by different data resources. To unify the building geometries, this study is aim to extract the spatial information from UAV-borne photos and the high-altitude aerial images to reconstruct 3D building models.
In order to extract the building geometries, techniques of Structure-from-Motion (SfM) and photogrammetry are implemented to acquire 3D information. SfM is to build up the relative orientations (ROs) among overlapped images of UAV photos. Once the exterior orientation (EOPs) of a UAV photo is established, a linear approach can be used to find the EOPs for other images by ROs. Regarding to the high-altitude aerial images, GPS and IMU are equipped to acquire EOPs of each image. An image pair including a UAV photo and a high-altitude aerial image can then be selected to find the features such as vertices and developed into building geometrics. With the essential spatial information, 3D building models can be reconstructed in computer environments for visualization and further applications. This study provides an integrated strategy to reconstruct the 3D building models by consistent geometries and improves the efficiency of spatial data acquisition and processing.

## 1. INTRODUCTION

With a variety of platforms to collect spatial information, techniques and algorithms of data processing have been prosperous developed. For instance, light detection and ranging (LiDAR) can generate point clouds of the real world and related processing methods have been proposed to build 3D models. Another example is the mobile mapping system (MMS), which provides a high motivation and great capability to acquire detailed spatial information and street-based 3D mapping (Hunter et al., 2006). Unmanned Aerial System (UAS/UAV) is a low-cost and has domain the developments of close-range photogrammetry for years. For high-altitude airborne images, a lot of algorithms have been proposed to extract the information in the 3D space from images. With the help of direct georeferencing, it is easier to determine the camera's location through GPS and IMU systems for precise data processing. A characteristic of both UAV and high-altitude aircraft is able to take oblique photography, which increases the ability of acquiring more spatial information from overlapped images.

To recognize the geometries and physical information is an inevitable work to transform 2D images into 3D spatial data (Zhang et al., 2008). Digital photogrammetry is sensitive in terms of camera pose, movements and geometry while high-accuracy 3D information can be achieved. Based on the collinearity condition, light tracking is used to build the relation among the perspective center, image points and objects. Techniques such as space resection, intersection and bundle adjustment are often used. On the other hand, computer vision emphasizes more on visualizing the 3D spatial data and real-time display. Fast data processing is one of features in computer vision that optimization is not even required. Comparing to non-linear data processing in photogrammetry, there are more linear transformations in computer vision.

In order to recovery the 3D models in virtual environments, a lot of useful methods and algorithms are available like extracting boundaries from point clouds and measuring the geometries of a building. As utilizing images to reconstruct 3D building models, the core techniques in this study are divided into computer vision and digital

photogrammetry. For data processing, there are several differences between computer vision and photogrammetry. In this research, computer vision provides a fast computation to connect a series of overlapped images through extracted features even if the knowledge of geometry is insufficient (Mayer et al., 2003). In addition, the extracted and processed spatial information is in the relative coordinate system. On the other hand, photogrammetry emphasizes the geometric accuracy of objects and a lot of robust techniques are adopted. There are merits and demerits of the proposed techniques such as efficiency and high accuracy. Therefore, this study tries to take appropriate approaches to reconstruct 3D models from UAV photos and high-altitude airborne imagery.

## 2. METHODOLOGY

In order to acquire the 3D spatial information and reconstruct 3D building models, computer vision and digital photogrammetry techniques are sued to get the essential data. The used data includes UAV photos without GPS and IMU information, ground-surveyed control points (GCPs), high-altitude airborne imagery by Leica RCD 30 and their EOPs. To achieve the purpose of reconstructing 3D building models, the data processing is divided into EOPs estimation of UAV photos and extracting building geometries from the image datasets as displaying in Figure 1.

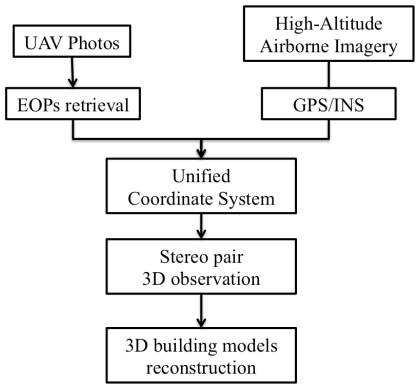



Figure 1. A proposed workflow of 3D building models reconstruction

### 2.1 Space Resection Based on Barycentric Coordinate System

In conventional photogrammetry, utilizing collinearity condition equations to find the EOPs of a camera is a widely used method for a vertical or near-vertical image. With 3 or more GCPs, least squares can be performed to get the optimized EOPs of a camera through single image space resection. However, for oblique or off-nadir aerial images (Figure 2), it would be tough to use collinearity condition equations to EOPs of an image due to the inappropriate initial values. Therefore, a space resection method based on barycentric coordinate without iterations (Li et al., 2015) was proposed to cope with this obstacle. This method has been proven useful in solving the EOPs of any rotation angles and off-nadir aerial images without initial approximations.

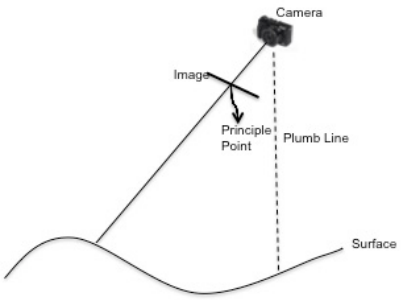



Figure 2. The relation between the camera and off-nadir image

To perform this method, a three-dimensional barycentry coordinate system uses a tetrahedron and its 4 vertices from $P_1$ to $P_4$ in the 3D space to express points as shown in Figure 3. A point $K$ $(X_K, Y_K, Z_K)^T$ in this system can be expressed as the combination of ration $\lambda_{Ki}$ and the vertices $P_i$ such that any point in this system is only related to the vertices as in Equation (1). Within the equation, the object coordinate system is represented by $W$ and the image coordinate is shown by $C$ with $i = 1, 2 \ldots n$ for any point. This coordinate system keeps the consistencies of points both in image and object coordinate systems. For EOPs estimation, GCPs are used to build up the transformation between the above spaces and coordinate transformation is help for EOPs retrieval.

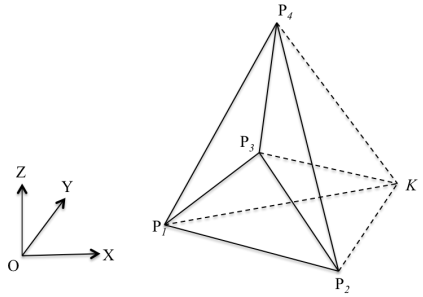


Figure 3. A barycentric coordinate system in 3D space

$$\begin{bmatrix} X_{C_i} \\ Y_{C_i} \\ Z_{C_i} \end{bmatrix} = \lambda_{i1} \begin{bmatrix} X_{P_{C1}} \\ Y_{P_{C1}} \\ Z_{P_{C1}} \end{bmatrix} + \lambda_{i2} \begin{bmatrix} X_{P_{C2}} \\ Y_{P_{C2}} \\ Z_{P_{C2}} \end{bmatrix} + \lambda_{i3} \begin{bmatrix} X_{P_{C3}} \\ Y_{P_{C3}} \\ Z_{P_{C3}} \end{bmatrix} + \lambda_{i4} \begin{bmatrix} X_{P_{C4}} \\ Y_{P_{C4}} \\ Z_{P_{C4}} \end{bmatrix} \tag{1}$$

By definition, if the origin $O\ (0,0,0)^T$ is selected as $P_1$, unit length of each axis can be written as $X'\ (1,0,0)^T$, $Y'\ (0,1,0)^T$ and $Z\ (0,0,1)^T$ corresponding to $P_2$, $P_3$ and $P_4$ respectively. Therefore, the ration $\lambda_{Ki}$ of point $K$ has a characteristic that $(\lambda_{K1}, \lambda_{K2}, \lambda_{K3}, \lambda_{K4}) = (X_K, Y_K, Z_K, 1 - X_K - Y_K - Z_K)$ in the normalized barycentric coordinate system while the summation of $\lambda_{K_{i=1\ldots4}} = 1$.

The conversion of GCPs from object space to image space is the foundation to find the EOPs of an image. Based on the light tracking, the relation of a camera, an image point and the ground object is displayed as Figure 4. An image point $a_i$ ($x_i$, $y_i$) where $i$ = 1, 2…$n$ in the 2D plane can be written as $a_i$ ($x_i$, $y_i$, $-f$) in the 3D image space with the camera focal length $f$. With the consideration of GCPs in the object space and image space through barycentric coordinate, the correlation is expressed as Equation (2) and the scale factor ($S$) is defined as Equation (3).

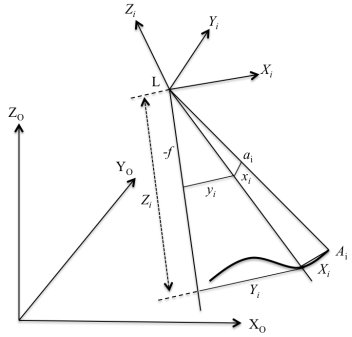


Figure 4. The collinearity condition between image space and object space

$$S \begin{bmatrix} x_{Ci} \\ y_{Ci} \\ -f \end{bmatrix} = \begin{bmatrix} X_{Ci} \\ Y_{Ci} \\ Z_{Ci} \end{bmatrix} = \sum_{j=1}^{4} \lambda_{ij} \begin{bmatrix} X_{P_{C_i}} \\ Y_{P_{C_i}} \\ Z_{P_{C_i}} \end{bmatrix} \tag{2}$$

$$\frac{X_{Ci}}{x_{Ci}} = \frac{Y_{Ci}}{y_{Ci}} = \frac{Z_{Ci}}{-f} = S \tag{3}$$

In the above equations, the purpose is to find the 4 vertices of $P_C$ such that there are 12 unknowns to be solved. From the expanded Equation (2), a single image point can provide 2 linear equations as shown in Equation (4) that at least 6 GCPs are necessary for least square solution. As long as the 4 vertices are decided, the locations of $GCP_i$ ($Xc_i$, $Yc_i$, $Zc_i$) in the image space can be acquired from Equation (1). With ground control points both in image and object spaces, 3D conformal transformation is introduced to build up the relations between these 2 coordinate systems by 7 elements (1 scale factor, 3 translations and 3 rotation angles). Among the parameters, the 3 translations and 3 rotation angles are referring to EOPs ($X^L$, $Y^L$, $Z^L$, $\omega$, $\varphi$, $\kappa$) of the camera.

$$\begin{aligned} f\left(\lambda_{i1}X_{P_{C1}} + \lambda_{i2}X_{P_{C2}} + \lambda_{i3}X_{P_{C3}} + \lambda_{i4}X_{P_{C4}}\right) + x_i\left(\lambda_{i1}Z_{P_{C1}} + \lambda_{i2}Z_{P_{C2}} + \lambda_{i3}Z_{P_{C3}} + \lambda_{i4}Z_{P_{C4}}\right) = 0 \\ f\left(\lambda_{i1}Y_{P_{C1}} + \lambda_{i2}Y_{P_{C2}} + \lambda_{i3}Y_{P_{C3}} + \lambda_{i4}Y_{P_{C4}}\right) + y_i\left(\lambda_{i1}Z_{P_{C1}} + \lambda_{i2}Z_{P_{C2}} + \lambda_{i3}Z_{P_{C3}} + \lambda_{i4}Z_{P_{C4}}\right) = 0 \end{aligned} \tag{4}$$

**2.2 Relative Orientations from Structure-from-Motion**

Structure-from-Motion is a powerful and efficient method for 3D model reconstruction. Related techniques of the algorithms had been proposed to process images such as scale-invariant feature transformation (SIFT) matching and essential matrix establishment. Among them, relative orientations between an image pair can be retrieved through conjugate features when a calibrated camera is used. The essential matrix is based on the epipolar geometry and correspondence of 2 image planes is constructed, containing a rotation matrix ($R$) and a translation vector ($T$) as shown in Figure 5. Such an essential matrix provides a mathematical model of a stereo pair in Equation (5), where $p_l$ and $p_r$ refers to the left and right image respectively and $E$ is the essential matrix. Due to the limitations of GCPs distribution and off-nadir UAV photos, it is a tough work to recover EOPs of each single image by space resection. Consequently, the retrieval of EOPs by a linear approach by ROs (He et al., 2014) is adapted in this study.

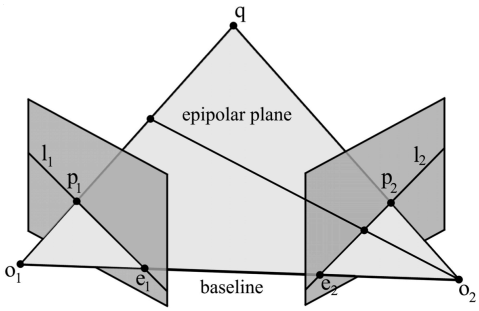


Figure 5. Epipolar geometry of a stereo pair

$$p_l E p_r = 0 \tag{5}$$

In order to estimate the rotation matrix and the translation vector, the proposed singular value decomposition (SVD) is implemented to acquire the rotation and translation by either 5-points algorithm (Nister, 2004) or 8-points algorithm (Hartley, 1997). As there are 6 RO elements of a stereo pair, the 8-points algorithm is elected and least square is able to optimize the ROPs. Practically, there are 4 possible solutions by the combination of 2 rotation matrixes and 2 translation vectors. To identify the correct $R$ and $T$, a well-known co-planarity model (Mikhail et al., 2001) is adapted to determine the most appropriate solution of the relative orientations. In Equation (6), the rotation matrix $R_r^l$ stands for the rotation from the right image to the left and RO elements are represented by a 3-by-4 matrix. By setting a 3-by-3 identical rotation matrix and a 3-by-1 zero translation vector to the left image, the ROPs of the right image can be acquired as Figure 6.

$$p_l^T \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix} R_r^l p_r = 0 \tag{6}$$

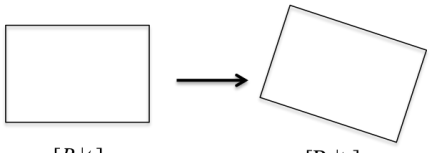


Figure 6. Relative Orientations between 2 stereo-images

During the process of SVD, extracted features from a stereo-image are normalized into 3D image space. The translation vector with three components in the 3D image space can be repressed as the base line between an image pair with scale ambiguity. In other words, the translation vector cannot appropriately describe the base line in the object coordinate. Therefore, a scale factor is necessary to recover the real translations in the object coordinate. Since the square summation of the 3 components is 1 and the unit (mm, cm or m) can be assigned according to the used coordinate system. A straightforward method is to use the base line in the object coordinate as the scale factor and recover the real translations.

While in this research, the EOPs are unknown and which leads difficulty to find the length of real base line. A substitute method to find the scale is to use the matched features in the 3D homogenous image space. In this stage, the interior orientations (IOs) can form a 3-by-3 IO matrix and map the 2D image points into 3D space. Because the scale factor of a translation vector is a constant respecting to the three components, this research tries to estimate the base line as the scale factor ($B$) by Equation (7). With 2 datasets of 3D image features, three-dimensional conformal

transformation (7 parameters transformation) can be performed to find the scale factor. The estimated scale factor is defined as Equation (8) via de-normalization and $S_i$ represents the scale of the 3D conformal transformation and $H$ is the height of the image with known EOPs. This equation is derived form base height ration (B/H) and the sign reflects the relative image scale. Therefore, EOPs of a series of UAV photos can be approximately estimated by ROs if the EOP of a UAV photo is found using space resection based on barycentric coordinate.

$$[T_X \quad T_Y \quad T_Z]^T = B[t_x \quad t_y \quad t_z]^T \tag{7}$$

$$B = (S_i - 1)H \tag{8}$$

**2.3 3D spatial information extraction**

In photogrammetry, the 3D spatial information can be acquired through stereoscopic measurement. As shown in Figure 7, 3D objects can be observed through a stereo pair with parallel eye base. From 2 stereo-images, the sizes of a 3D object can be determined by the fundamental concept of 3D observation (Jensen, 2006). Practically, this can be achieved via collinearity condition that the same objects appear in more than one image.

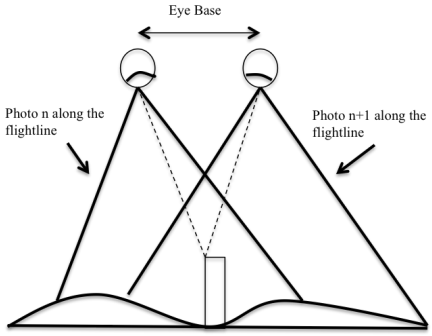



Figure 7. Stereoscopic viewing and 3D object observation

The well-known linearized collinearity condition equations are used to extract 3D spatial information, and the robust estimation is implemented. In this optimization, initial approximations of object coordinates have to be provided for iterations. However, this study tries to recover the 3D information from a UAV photo and a high-altitude aerial-borne image. Conventional technique of using air base, parallax, image-coordinates and camera IO to estimate the initial 3D object coordinates would be tough in this situation. An alternative method of direct georeferencing using a single image with known EOPs (Schward et al., 1993) had been proposed to estimate the initial approximations as Figure 8 and Equation (9).

In Equation (9), $(X^O, Y^O, Z^O)^T$ describes the estimated initial values of an object in the real coordinate system. Within the right side, $(X^L, Y^L, Z^L)^T$ indicates the position of the camera, $s_i^O$ represents the scale of the image and is computed as $s_i^O = -f / Z^L$ and $(x_i, y_i, -f)$ is an image point where $f$ is the focal length. Thus, initial approximations of an object in real world coordinate can be achieved by a single image with its known EOPs.

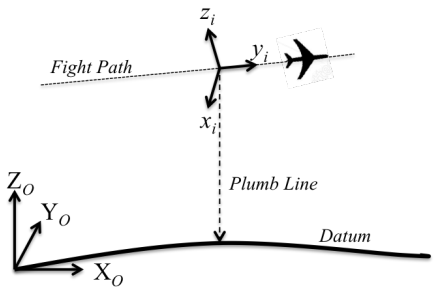
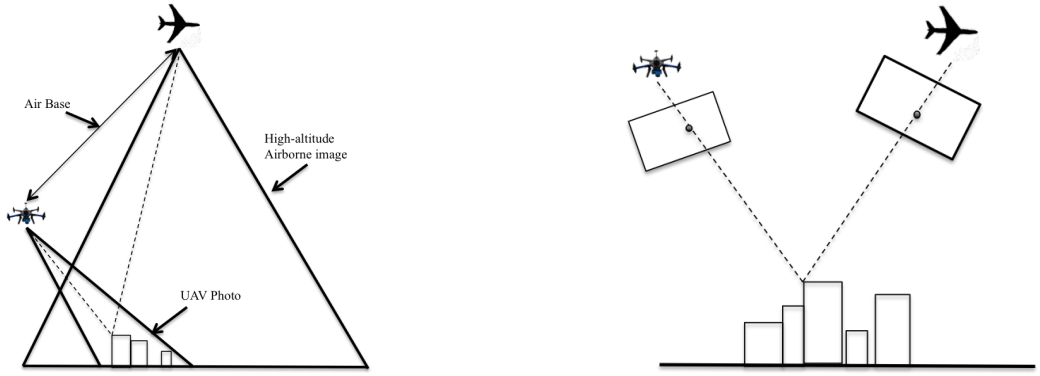



Figure 8. The relation between the image frame and the object space frame

$$\begin{bmatrix} X^O \\ Y^O \\ Z^O \end{bmatrix} = \begin{bmatrix} X^L \\ Y^L \\ Z^L \end{bmatrix} - s_i^o R_i^O \begin{bmatrix} x_i \\ y_i \\ -f \end{bmatrix} \tag{9}$$

The space relation of a UAV photo and a high-altitude airborne image is displayed as Figure 9(1). A robust

optimization of forward space intersection is adapted to acquire the 3D spatial information. The 3D objects extraction is constrained by Equation (10) by a UAV photo and a high-altitude airborne image. With essential 3D building geometries, the boundary representation (b-rep) (Jackson et al, 1995) and the physical information are assist to reconstruct building models in virtual systems. In order to keep the eye base parallel for 3D observation, this study gives the weights based on altitude ratio and elements of UAV photos can be virtually processed.




(1) (2)

Figure 9. A UAV photo and a high-altitude airborne image

$$\frac{X^O - X_U^L}{x_u - X_U^L} = \frac{Y^O - Y_U^L}{y_u - Y_U^L} = \frac{Z^O - Z_U^L}{z_u - Z_U^L}$$
$$\frac{X^O - X_A^L}{x_a - X_A^L} = \frac{Y^O - Y_A^L}{y_a - Y_A^L} = \frac{Z^O - Z_A^L}{z_a - Z_A^L} \quad (10)$$

## 3. RESULTS AND DISCUSSIONS

### 3.1 Validation of Space Resection based on Barysentric Coordinate

To ensure the applicability of this method to acquire exterior orientations, the measured positions and rotations of the high-altitude airborne images are used to exam the feasibility. The given EOPs had been rectified by bore-sight and lever-arm calibrations and will be compared to the EOPs from space resection. In the test image, 6 GCPs are selected and using the proposed space resection based on barycentric coordinate. To get reliable EOPs, ground control points have better distributed averagely in the whole image, otherwise the quality of the results would be unstable. Table 1 shows the comparisons of the measured and the estimated EOPs of an aerial image and it proves that the results are close to measured EOPs.

One of the merits of this space resection is that it can cope with any tilt angle even to off-nadir images. However, there is a disadvantage that it is unable to detect the blunders or mistakes comparing to robust space resection. In addition, ground control points are manually measured in an UAV photo and inevitable errors may occur. This may lead to errors when computing EOPs as in the example that there is an obvious difference of $Y^L$.

Table 1. Exterior orientations measured by GPS/INS and estimated from GCPs

| EOP / Method | $X^L$ (m) | $Y^L$ (m) | $Z^L$ (m) | $\omega$ (degree) | $\varphi$ (degree) | $\kappa$ (degree) |
|---|---|---|---|---|---|---|
| GPS/INS | -31401.096 | -53721.980 | 570.778 | 0.03313 | 34.872 | 89.669 |
| Space Resection | -31400.824 | -53717.706 | 569.035 | -0.40326 | 35.015 | 89.799 |
| Absolute Difference | 0.272 | 4.274 | 1.743 | 0.436 | 0.143 | 0.130 |

### 3.2 UAV Photos and High-Altitude Airborne Images

Two datasets are used in this research, including UAV photos taken by Canon EOS-M22 with fixed focal length and Leica RCD30 Oblique airborne scanner. Leica RCD30 oblique airborne scanner is equipped with 5 cameras containing nadir, forward, backward, right and left sensors. Both of the used platforms are capable of taking tilt images, which increase the ability of acquiring more spatial information. The coordinate system is JGD2000/Zone 9 both in ground survey control points and the GPS/INS of Leica RCD30 oblique airborne scanner. In addition, the interior parameters had been calibrated and image lens distortion is corrected.

Figure 10 and Figure 11 are the flying paths of UAV and high-altitude aircraft. As for UAV, it only scanned along the pre-planned path without multi-scans or took photos around a full view of a specific building. Therefor, spatial information extracted from UAV images may be insufficient for 3D building geometries especially when there are occlusions. Although there are a lot of datasets from both aircrafts, it is necessary to select appropriate images so that 3D building geometries and physical information can be identified.

Figure 10. The fly path of unmanned aerial vehicle

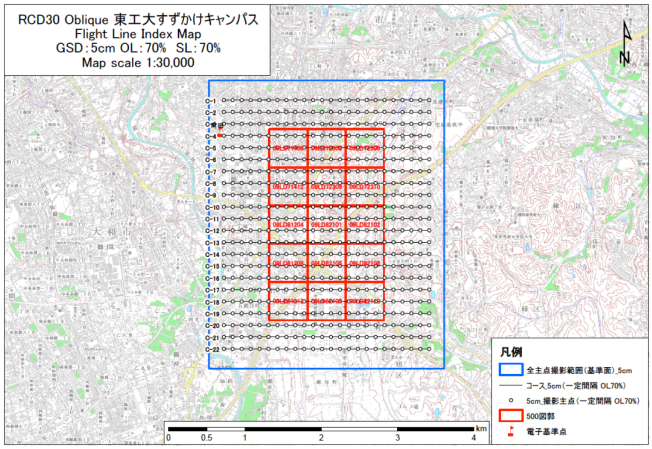
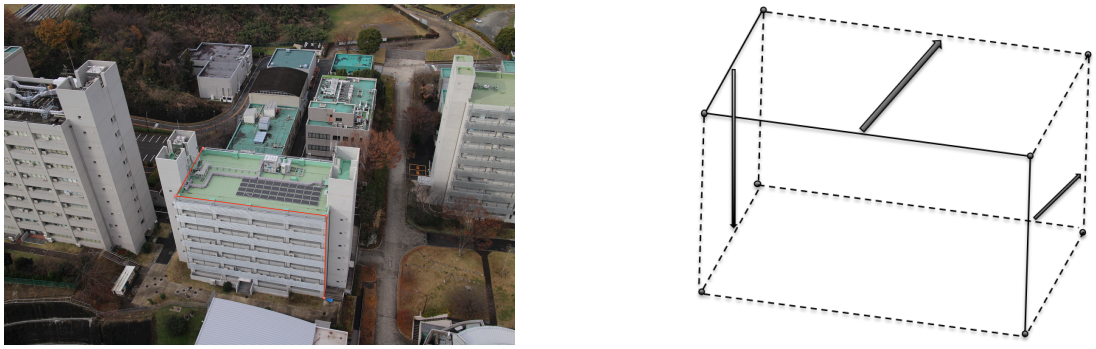



Figure 11. A planned fly path of the high-altitude aircraft

### 3.3 Reconstruction of 3D Building Model and Discussion

In this study, LOD2 building block models specified by OGC CityGML will be reconstructed and a sample result is displayed and analyzed. Figure 12 (1) shows a simple building for reconstruction and 3D information of those marked vertices are extracted. For a regular building, the shape is similar to a rectangle or a square so that boundaries can be estimated through physical observations. To recover a complete 3D building model, symmetric structures of a regular building can provide additional information even though there are occlusions. With sufficient points and line-based information, the boundaries of a building can be depicted by b-rep and occlusions may be conquered shown as Figure 12 (2).



(1) (2)

Figure 12. Boundary extraction and 3D building block model reconstruction

Within the processes of extracting 3D object information, three datasets of stereo image are used to compare the results and 2D construction plan is to validate the sizes. The comparisons of using 2-image-stereo contain UAV stereo-pair, high-altitude airborne stereo-images and their combined stereo-pair. Essential 3D building information includes three elements of length, width and height. Usually, it is better to collect images around a building in order to acquire reliable spatial data and avoid occlusions. However, the UAV photo datasets only present specific perspectives of a building and it is hard to get an overall validation. An example of the reconstructed 3D building block model is shown in Figure 13 as a LoD2 representation. The 3D sizes are compared in Table 2 with the differences to the reference data followed by the computed results.

Table 2. Comparisons of 3D building sizes using different stereo pairs

| Stereo pair / Size (m) | High-altitude Airborne Image | UAV Photo | UAV and High-altitude Airborne Image | Reference From 2D Construction Plan |
|---|---|---|---|---|
| Length | 30.326 (1.526) | 29.180 (0.380) | 28.658 (0.142) | 28.800 |
| Width | 16.381(1.381) | 17.242 (2.242) | 14.801(0.199) | 15.000 |
| Height | 24.205 (1.705) | 20.237 (2.263) | 20.510 (1.990) | 22.500 |
| EOPs Resource | Measured by GPS/INS | Estimated by GCPs | UAV- Estimation Airborne- GPS/INS | - |

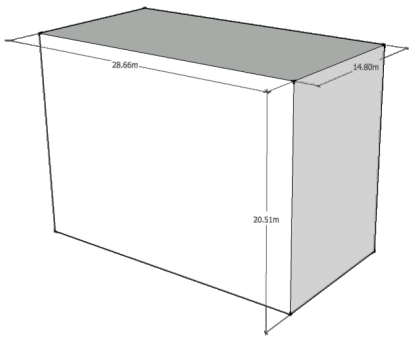



Figure 13. An example of the reconstructed 3D building model using UAV and high-altitude airborne stereo image

Since the UAV is not equipped with GPS/INS and GCPs are manually extracted from images so that there may be inevitable errors leading to inaccurate results. Moreover, ROs of the estimated relative orientation by essential matrix are not accurate enough. If high accuracy is required, EOPs of UAV photos acquired from ROs have better double adjusted if there are GCPs in the photo. If the EOP of UAV stereo pair is not accurate and errors from image matching will lead to incorrect 3D information. From the absolute difference values between the constructed results and the reference data, the reliability and accuracy is improved with the help of high-altitude airborne imagery.

The third result in Table 1 combing different datasets includes 2 UAV photos and 1 high-altitude airborne image in order to improve the accuracy. Consequently, a similar to SfM multi-view stereo (MVS) technique with only building corner extraction is launched by robust space intersection. For SfM-MVS consideration, more-overlapped images can increase the reliability of acquiring 3D information. In addition, assigning different weight to the used

image datasets is helpful when performing photogrammetry techniques if there are several datasets. When the fundamental 3D building block model is established and related works cab be expected to do such as texture mapping or visualization in computer systems.

## 4. CONCLUSION

With a variety of data resources, to extract and integrate spatial information is still developing. Since close-range photogrammetry has become a popular research topic, more spatial information can be retrieved for applications. But on another perspective, there are more challenges to process those images in order to extract desired spatial information. The proposed idea in this study of integrating UAV photos and high-altitude airborne imagery for 3D building models reconstruction combines photogrammetry and computer vision. In the experiments, EOPs of UAV photos can be effectively acquired by space resection and estimated by relative orientations. As the datasets are in the same coordinate system, the integration of UAV photos and high-altitude airborne imagery can be implemented to extract 3D building information.

The 3D information is extracted based on the stereoscopic and three results are compared in this paper. The results indicate the feasibility of using different image datasets for space intersection. By comparing the results and the references, it is proved that this method is applicable. Moreover, the robust space intersection is able to balance the errors from measured and calculated EOPs. In the example, there are two UAV photos and a high altitude airborne imagery that increase the constraints of light tracking. If more images are provided to retrieve 3D spatial information, the accuracy will be improved. The approach of merging photogrammetry and computer vision therefore can be a novel strategy to extract 3D information from images in the near future.

## ACKNOWLEDEMENT

The authors would like to thank Yoshihiro Ishitsuka (Pasco Corporation) for data confirmation.

## REFERENCES

Hartley, R., 1997. In defence of the 8-point algorithm. Proceedings of IEEE International Conference on Computer Vision 19 (6), pp. 1064-1070.

Hunter, G., Cox, C. and Kremer, J., 2006. Development of a commercial laser scanning mobile mapping system -streetmapper. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 36.

He, F. and Habib, A., 2014. Linear approach for initial recovery of the exterior orientation parameters of randomly captured images by low-cost mobile mapping systems. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol.XL-1, pp.149-154.

Jackson, D., 1995. Boundary representation modelling with local tolerances. Proceedings of the Third Symposium on Solid Modeling and Applications, pp. 247-254

Jensen, J.R., 2006. Remote sensing of the environment: an earth resource perspective, edited by Clarke, K.C, Pearson, pp. 163-169.

Li, J., Hu, Q. and Ai, M., 2015. A non-iterative space resection method based on barycentric coordinates. Acts Geodaetica et Cartographia Sinica, 44 (9), pp. 988-994.

Mikhail, E.M., Bethel, J.S. and Glone, J.C., 2001. Introduction to modern photogrammetry, John Wiley & Sons, New York.

Mayer, H., Mosch, M. and Peipe, J., 2003. Comparison of photogrammetric and computer vision techniques: 3D reconstruction and visualization of Wartburg Castle. Proceedings of the XIXth International Symposium, CIPA, pp. 103-108.

Nister, D., 2004. An efficient solution to the five-point relative pose problem. Pattern Anal. Mach. Intell. IEEE Trans, 26 (6), pp. 756-770.

Schward, K.P., Chapman, M.A., Cannon, M.W. and Gong, P., 1993. An integrated INS/GPS approach to the georeferencing of remotely sensed data. Photogrammetric Engineering & Remote Sensing, 59(11), pp. 1667-1674.

Zhang, C. and Yao, W., 2008. The comparisons of 3D analysis between photogrammetry and computer vision. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Science, Beijing, 58, Part B3b, pp. 1-4.