

THE COMPARISONS OF 3D ANALYSIS BETWEEN PHOTOGRAMMETRY AND COMPUTER VISION

Chunsen ZHANG, Wanqiang YAO^a

Dept. of Survey Xi'an University of Science & Technology,
No.58 Yantazhonglu, Xi'an, 710054 China -zhchunsen@sina.com

Commission III WG III/1

KEY WORDS: Digital photogrammetry, Computer vision, 3D Analysis Method

ABSTRACT:

Photogrammetry and computer vision are coherent in research content, but since they are come from different subject so there are great differences in researching method and application. The paper reviews the space resection, relative orientation and absolute orientation which is the kernel content of Photogrammetry 3D space computational, and comparisons it with the computer vision. As while as, explain its strongpoint and its shortage and given an example of the combine of the two methods. It can be seen from this comparing that the most theory and algorithm of computer vision are great value to digital photogrammetry, and the same that photogrammetry adjustment computation method has also important meaning to improve the measure precision of computer vision.

1. INTRODUCTION

Digital photogrammetry is the theory and method based on the keystone of digital image and photogrammetry, the computer technique, digital image processing, computer vision, pattern recognition and other subject. It is a branch of photogrammetry subject that pick-up geometry and physics information of taking object from the image, and it expressed by the way of digital[Zhang zhuxun 1996]. The aim of computer vision research is make the computer have the ability of cognize 3D scene information from 2D image. It includes the shape of geometry, position, movement, pose and so on, and describe, storage, identifying and comprehending to them[Ma Songde 1998]. The common ground between digital photogrammetry and computer vision is to obtain 3D information from images. Because digital photogrammetry has developed from traditional photogrammetry, its purpose of 3D analysis method is based on survey relief map, which fit for aviation photography condition, and from the point of survey view. The 3D analysis method of computer vision is based on the theory of vision analysis that founded by Marr. According to the theory frame, describing vision information can be divided into three phases which from two dimension image information originally to three dimension scene expression. In the first phases, the original image is processed, extract feature such as the corner points, edges, texture, lines and boundary. The "essential map" or "basic map" is constituted; in the second phases, rebuilding the taken objects' deepness of visible part, normal direction, and figure by input image and basic map in the coordinate system as observer centre. These contain deepness information, but it is not real 3D expression. It is named 2.5 dimensions expression. The phases, which rebuilding, expression and identify the taken 3D objects by input image, basic map and 2.5 dimensions map in the coordinate system as object centre, is the third phases. Although there is shortage in the theory of Marr, It is confessed that Marr theory discovers the possibility and basic method to rebuilt 3D object form 2D image by the numbers. That it to say, the 3D analysis method of digital photogrammetry is aim at the purpose of survey and mapping, together with terrain expression. The 3D analysis method of digital of computer

vision is aim at the more common situation. but so far, the research of computer vision mostly have still rested on lab, the practicality system comparatively less.

Along with the developed of the two subjects, it is more and more patency for them to tend towards consistent in the way of 3D analysis. From the point of digital photogrammetry view, it has being from traditional photogrammetry which give priority to terrain survey to expand digital close photogrammetry that the way of photograph is more complex than aviation photograph. It is including industry parts high precision measurement, rebuilt city 3D buildings and so on application area. For all these season, it is request photogrammetry field to develop more run-of-mill 3D analysis theory. From the point of computer vision view, it should walk out lab to give up ideal model study. Under certain condition, the research aim should be advanced from 3D reconstruct to high precision measurement or measurable 3D reconstruct, until tend towards practicability.

The purpose of the paper is comparisons digital photogrammetry and computer vision key process tache in 3D analysis and searching the points which can use for reference each other between them. The biggest difference is resection in space, relative orientation and absolute orientation in each key tache, so comparisons computer vision and digital photogrammetry's difference in above content. A example that combine each another is given.

2. THE DIFFERENCES BETWEEN PHOTOGRAMMETRY AND COMPUTER VISION

2.1 Space resection

The mathematical representation which indicates that the space $A(X, Y, Z)$, and its corresponding image points $a(x, y)$ lie on the same straight line passing through the projection centre $S (X_s, Y_s, Z_s)$ is as following:

$$\begin{bmatrix} x - x_0 \\ y - y_0 \\ -f \end{bmatrix} = \begin{bmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{bmatrix} \begin{bmatrix} X - X_S \\ Y - Y_S \\ Z - Z_S \end{bmatrix} \quad (1)$$

From this can get the most important equations in photogrammetry—collinearity condition equation as following

$$\begin{aligned} x - x_0 &= -f \frac{a_1 \cdot (X - X_S) + b_1 \cdot (Y - Y_S) + c_1 \cdot (Z - Z_S)}{a_3 \cdot (X - X_S) + b_3 \cdot (Y - Y_S) + c_3 \cdot (Z - Z_S)} \\ y - y_0 &= -f \frac{a_2 \cdot (X - X_S) + b_2 \cdot (Y - Y_S) + c_2 \cdot (Z - Z_S)}{a_3 \cdot (X - X_S) + b_3 \cdot (Y - Y_S) + c_3 \cdot (Z - Z_S)} \end{aligned} \quad (2)$$

Where (x, y) is the coordinate of the image point, (x_0, y_0) is the coordinate of the image point is the coordinate of the principal point, f is the principal distance. The values x_0, y_0, f are referred to as the elements of interior orientation of the image. (X, Y, Z) is the 3D object-space point coordinates, (X_S, Y_S, Z_S) is the object-space coordinate of projection centre S. $(a_i, b_i, c_i, i = 3)$ are the functions of the 9 direction cosines which constituent by 3 exterior orientations angular elements.

Above value have the definitude physics meaning. The elements of interior orientation and the elements of exterior orientation are separated in the equation.

Photogrammetry space resection is the process of photogrammetry that use some known object space point(control point),its related point in image and the collinearity equation to calculate the elements of interior orientation, exterior orientation and other additional parameter. From equation (2) it can know that every point can give two collinearity equations. Generally it takes more than six control point to do the adjustment computation with the least squares method in photogrammetry space resection.

Generally photogrammetry use metric camera, so it use the equation(2). It is the main task of photogrammetry to determine the six elements of exterior orientation of the image.

Computer vision often describe the same relation in homogeneous coordinates projection equation as following

$$Z_c \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & x_0 & 0 \\ 0 & f_y & y_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & t \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = M_1 M_2 \bar{X} = M \bar{X} \quad (3)$$

Where f_x, f_y is equivalent focus in X and Y direction parting. $\bar{X} = (X, Y, Z, 1)^T$ is space points' homogeneous coordinates under object-space coordinate system. Z_c is space points' depth under camera coordinate system. M is a 3×4 projection matrix. M_1 is interior orientations element matrix. M_2 is exterior orientations element matrix. The advantage of this is that it can be expression by a linear equation simplify.

Contrast with photogrammetry, in computer vision, the process which obtain interior orientations element of image intituled camera calibration. Because non-metric camera is widely used in computer vision field, the research of camera calibration technique with high precision is very importance work. In order to achieve the purpose, combining the theory of photogrammetry and computer vision, many scholars have studied the camera calibration technique thoroughly. And many methods have been brought forward.

Comparing with the collinearity equation in photogrammetry described before, projection matrix in computer vision is more compact in math expression form. And because the matrix determine the relationship between the space point's coordinate and its image point coordinate, so it can go without

calculate the camera's elements of interior orientation and elements of exterior orientation. That is to say the matrix M itself is the camera parameters, but those parameters are not have specifically physical meaning as the principal distance, principal point in photogrammetry, it is just a integrate expression of camera parameters.

Along with the digital camera popularization and digital close photogrammetry is used widely. It has been interest issue for photogrammetry operator to research the camera calibration technique. As while as, the method of least squares used by photogrammetry operator also provide the ample idea for computer vision operator.

2.2 Relative orientation

The relative orientation in photogrammetry is the project rays pass corresponding image points meet the condition that intersect each other. that is the geometric coplanar condition photogrammetry space based line and corresponding image rays are coplanar.

$$F = \begin{vmatrix} b_x & b_y & b_z \\ u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \end{vmatrix} = 0 \quad (4)$$

Where

$$\begin{bmatrix} u_1 \\ v_1 \\ w_1 \end{bmatrix} = R_{left} \begin{bmatrix} x_1 \\ y_1 \\ -f \end{bmatrix} \quad \begin{bmatrix} u_2 \\ v_2 \\ w_2 \end{bmatrix} = R_{right} \begin{bmatrix} x_2 \\ y_2 \\ -f \end{bmatrix} \quad (5)$$

R_{left} and R_{right} is left and right image' orientation matrix separately. $(x_1, y_1), (x_2, y_2)$ is left and right image' coordinate of homonymy points.

For the 5 homologous image points, there may be 5 simultaneous equations. Since the first order terms of Taylor expansion are taken in these equations, approximations should be performed to correct the values of the coefficients and the constants successively until the desired accuracy has been achieved. If there are observations from more then 5 homologous image points, least squares adjustment should be applied.

Since the intersection point set of homologous image points constitutes the digital terrain model (DTM), photogrammetry relative orientation is used establish DTM.

In computer stereo vision, the relative orientation not only uses in photograph geometric three dimensional rebuilding but also in stereo vision camera calibration. Computer vision often calculate the basic matrix F and the relationship between the two camera(two images)from two camera matrix $M_1 M_2$,that is the expression form of relative resection in computer vision with as following:

$$u_2^T F u_1 = 0 \quad (6)$$

where: u_1, u_2 separately are the image homogeneous coordinates of corresponding image points, $F = [m]_{\times} M_2 M_1^{-1}$, $[m]_{\times}$ is the skew-symmetric matrix defined by $m = m_2 - M_2 M_1^{-1} m_1$, $m_1 m_2$ are separately the three dimensional shift of the two camera. Matrix F is an important matrix in computer video vision and computer movement vision, it was called basic matrix. The hot problem researched in computer vision is how to calculate the basic matrix.

When given 8 points (or more) correspondences across two images, the F can be find from linear equations. The famous method of "8 points method" as follows.

Assuming $F_i' = (u_i', v_i', 1)$ is normalized vector (namely $u_i' = x_i'/f$, $v_i' = y_i'/f$) of left certain image point under the image space assistant coordinate system. $F = (u_i, v_i, 1)$ is normalized vector of right certain image point under the image space assistant coordinate system. $T = (B_x, B_y, B_z)$ is a vector of photograph base line under left camera image space assistant coordinate system. The vector coplanar equation for homologous image points is:

$$(F_i')^T \cdot (T \times R) F_i = 0 \quad (7)$$

Assume $E = (T \times R)$
Introduce skew-symmetric:

$$[T]_{\times} = \begin{bmatrix} 0 & -B_z & B_y \\ B_z & 0 & -B_x \\ -B_y & B_x & 0 \end{bmatrix}$$

The E can be express: $E = [T]_{\times} R$
The equation (7) can be express: $(F_i')^T E F_i = 0$ (8)
Assume

$$E = \begin{bmatrix} e_1 & e_2 & e_3 \\ e_4 & e_5 & e_6 \\ e_7 & e_8 & e_9 \end{bmatrix}$$

The equation (8) can be express in the form of linear equation:

$$\begin{bmatrix} u_1' u_1 & u_1' v_1 & u_1' & v_1' u_1 & v_1' v_1 & v_1' & u_1 & v_1 & 1 \\ u_2' u_2 & u_2' v_2 & u_2' & v_2' u_2 & v_2' v_2 & v_2' & u_2 & v_2 & 1 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ u_3' u_3 & u_3' v_3 & u_3' & v_3' u_3 & v_3' v_3 & v_3' & u_3 & v_3 & 1 \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \\ \dots \\ e_9 \end{bmatrix} = 0 \quad (9)$$

Equation (9) is a homogeneous equation. The solutions obtained contain a scale factor. After obtaining matrix E and SVD, the value of absolute rotation and shift vector, up to scale can be obtained.

Heipke [Heipke, 1997] point out: Comparison with the photogrammetry's method and 8 points method there are some merit. It does not need initial value of unknown. No need iterative calculates. But in order to use 8 points method, it is has strict request for the model points in the direction of depth distributing relative to projection centre distance. Because camera principal distance and aviation highs need to normalize when use 8 points method. The process is not suit for photogrammetry. Comparing with the method of photogrammetry, it is not enough steady in geometry when the method is used in aerophotogrammetry.

2.3 Absolute orientation

Absolute orientation of photogrammetry is the process of determining the scale and orientation element of geometry model which built by relative orientation. This kind of different coordinate system's transform is 3D space resemble transform in mathematics.

Assuming λ is the scale factor. a_i, b_i, c_i to be the direction cosines of the included angles between the corresponding axes of the model coordinate system and of the ground control system coordinate system. (U, V, W) is model coordinate and (X, Y, Z) is the photogrammetry coordinate of model points. $(\Delta X, \Delta Y, \Delta Z)$ is shift vault of the difference coordinate system origin.

When the 3D linear transformation is written in matrix form, it becomes:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \lambda \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{bmatrix} \begin{bmatrix} U \\ V \\ W \end{bmatrix} + \begin{bmatrix} \Delta X \\ \Delta Y \\ \Delta Z \end{bmatrix} \quad (10)$$

In the computer vision, this is similar to the absolute orientation of photogrammetry, there are some questions such as transforming every point that in camera coordinate system to the coordinate in the scene absolute coordinate system. In computer vision field usually adopt skew-symmetric matrix (quaternion) to calculate rigid transforming question between the two coordinate systems.

The characteristic of skew-symmetric matrix is that it inducts other parameter to express the rotation matrixes R, since R is an orthogonal matrix, according to definition, R can be express:

$$R = (I + S) (I - S)^{-1} \quad (11)$$

S is a skew-symmetric matrix. Thus $ST = -S$

$$S = \begin{bmatrix} 0 & -c & -b \\ c & 0 & -a \\ b & a & 0 \end{bmatrix} \quad (12)$$

Since absolute orientation can be express:

$$X_i = R X_i' + T \quad (13)$$

Assume: centre of gravity coordinate of \bar{X} and \bar{X}' is X_i and X_i' . From Eqs. (13) We have:

$$(X_i - \bar{X}) = R(X_i' - \bar{X}')$$

Thus

$$(X_i - \bar{X})^T = (X_i' - \bar{X}')^T R^T = (X_i' - \bar{X}')^T R^{-1} \\ = (X_i' - \bar{X}')^T (I - S)(I + S)^{-1}$$

$$\text{And } (X_i - \bar{X})^T (I + S) = (X_i' - \bar{X}')^T (I - S) \quad (14)$$

$$\text{Let: } (X_i - \bar{X})^T = A, (X_i' - \bar{X}')^T = B$$

$$\text{We have } \begin{aligned} A(I + S) &= B(I - S) \\ (A + B)S &= (B - A) \end{aligned} \quad (15)$$

According Eq. (15), the skew-symmetric matrix can be obtained. Form S, the rotation matrixes R can be constructed. Thus T can be calculated by centre of gravity coordinate.

It is clear: calculating the rotation matrixes R by the method of skew-symmetric matrix (quaternion) is that a linearity equation solution. It need not initial value.

3. CONCLUSION

From the previous compare we can see, the kernel content of photogrammetry 3D space analyzed are the space resection, relative orientation and absolute orientation which respective correspond to the imaging process express of pinhole camera model, geometric relation between stereo images and the absolute transform of coordinate systems. Moreover, computer vision use many photos to rebuild the three dimensional structure of the object which called three dimensional rebuilding problem. Namely, it is the classical space resection and method of intersection in photogrammetry. Interpolation corresponds to the creation of DEM in photogrammetry. Thereby, photogrammetry is very similar to computer vision in research content. Secondly, there is an obvious distinguished between computer vision and photogrammetry that is computer vision adopts more homogeneous coordinate to express different kinds of equations. In this way the expressive equation is a linear equation; via the equation transform the best result can be gained. Whereas photogrammetry often uses nonlinear equation to express the same question, and adopts the way of survey data processing—least squares adjustment which runs through the whole process to calculate. The process is linearization of nonlinear equation, measuring observation value, ensuring initial value and alternate calculate. The linearity method of computer vision usually accomplish by mathematic transform. So every unknown in equation has not strict physical meaning. But every unknown in photogrammetry equation generally has strict physical meaning. Although, it is obvious that the research method of computer vision is more compact than photogrammetry's, the calculational precision of computer vision is lower.

Along with the widely application of digital close photogrammetry, the obtainable method of image data is more and more similar to computer vision, but the classical photogrammetry is more and more unable to deal with these data. Thereby, some new research productions in the field of computer vision such as rebuilding theory, method as well as the new describable tools of obtaining image development which aims at non-metric camera. Undoubtedly all of these are worth to and use for reference and learn for photogrammetry.

REFERENCES

- Christian Heipke.1997. Automation of interior, relative, and absolute orientation. *ISPRS Journal of Photogrammetry and Remote Sensing*, 52
- Fabio Remondino.2002. Image Sequence Analysis for Human Body Reconstruction. *ISPRS Close-Range Imaging, Long-Range Vision*. Greece, 590-595
- Hartley, R. and Zisserman, A. (2000) *Multiple View Geometry in Computer Vision*. Cambridge University Press, UK.
- MA Song-de,ZHANG Zheng-you. Computer Vision: Fundamentals of Computational Theory and Algorithms [M]. Beijing: Sciences Press, 1998:151-173(in Chinese)
- ZHANG Zhuxun. *Digital Photogrammetry*. Wuhan university Press
- Project supported by the Education Foundation of Shaanxi Committee Education (No. 06KJ239)