УНИВЕРСИТЕТ ПО АРХИТЕКТУРА, СТРОИТЕЛСТВО И ГЕОДЕЗИЯ
Международна юбилейна научно-приложна конференция УАСГ2012

15-17 НОЕМВРИ 2012
15-17 NOVEMBER 2012

International Jubilee Conference UACEG2012: Science & Practice
UNIVERSITY OF ARCHITECTURE, CIVIL ENGINEERING AND GEODESY

PHOTOGRAMMETRIC INTERPRETATION OF EPIPOLAR GEOMETRY (EXTENDED)

B. Marinov[1]

Keywords: *Digital Photogrammetry, Stereoscopic Images, Computer Vision, Geometric Interpretation, Image Orientation*

Research area: *Photogrammetry and Remote Sensing*

ABSTRACT

The epipolar geometry is a modern method for description of relationships between corresponding images in stereo couple. It is based on usage of homogeneous coordinates for description of transformation of coordinates in the images.
The photogrammetric interpretation of epipolar relations is suggested. The calculations of fundamental and essential matrices are formulated in terms of photogrammetric equations in aerial and close-range photogrammetry. The matrices are presented in the form for obtaining the traditional parameters of photogrammetric transformations. The formulated transformations are adopted for calculating the parameters of orientation of stereo couples. The transformations could be used for digital 3D vectorization or for finding the corresponding matching points in stereo images.

1. Introduction

The idea of usage of information from stereo images without usage of three-dimensional metric data for calibration is introduced by Olivier Faugeras in 1992 [1]. The main assumption is that there is a stereo camera system that is capable, by comparing the two images, of establishing some correspondence between them. For such systems there are not known their intrinsic and extrinsic parameters, which is known as the uncalibrated system.

The usage of homogeneous coordinates for spatial processing of spatial information is presented in details by Olivier Faugeras [2]. The two viewpoints could be a stereo pair of images, or a temporal pair of images, which are described in terms of the geometry of two different images of the same scene, known as the epipolar geometry.

Maybank and Faugeras [4] showed that self-calibration can be computed from a single uncalibrated camera that undergoes some displacement. All that is assumed is that some point matches can be established between images. They developed an algorithm that

[1] B. Marinov, Dr., 1 Hristo Smirnenski boul., Sofia, Bulgaria, dimitrovborislav@abv.bg

requires the camera to undergo a minimum of three displacements; with a minimum of seven matched points between successive images. From this information it is possible to extract the inner parameters of the camera.

2. Basics of Epipolar Geometry

The mathematical principle of binocular vision is that of triangulation. For any single image, the three-dimensional location of any visible object point must lie on the straight line that passes through the centre of projection and the image of the object point. Determination of the intersection of two such lines generated from two independent images is called triangulation. The process of establishing such matches between points in a pair of images is called *correspondence* between matching points in two images.

2.1. Definitions of main terms in epipolar geometry

At first it might seem that correspondence requires a search through the whole image, but the *epipolar constraint* reduces this search to a single line. To see this, we consider Fig. 1.

- Epipole $\boldsymbol{e}_i$ is a cross of baseline with corresponding image plane. Epipole is a projection of an optical center of one image onto the plane of another image.
- Epipolar plane is a plane defined by any space point and two projection centers.
- Epipolar line is cross line between epipolar plane and image plane. All epipolar lines intersect in the epipole.
- All epipolar planes intersect in baseline.

2.2. Usage of homogeneous coordinates

The usage of homogeneous coordinates requires increasing the dimensions of vectors. The fourth coordinate is used in three-dimensional coordinate system.

$$\boldsymbol{G} = \begin{bmatrix} X & Y & Z & 1 \end{bmatrix}^t \quad \boldsymbol{g} = \begin{bmatrix} x & y & z & 1 \end{bmatrix}^t \tag{1}$$

where $\boldsymbol{G}$ and $\boldsymbol{g}$ are homogeneous spatial coordinates of vectors in object and image.

For two-dimensional coordinate system is added a third coordinate. Image points in planar homogeneous coordinates $\tilde{\boldsymbol{m}}_1$ и $\tilde{\boldsymbol{m}}_2$ are presented by the relations:

$$\tilde{\boldsymbol{m}}_1 = \begin{bmatrix} u_1 & v_1 & 1 \end{bmatrix}^t \qquad \tilde{\boldsymbol{m}}_2 = \begin{bmatrix} u_2 & v_2 & 1 \end{bmatrix}^t . \tag{2}$$

Such approach allows the processes of rotation and translation to be presented in general way by the multiplication of vector and matrix.

2.3. Fundamental and Essential matrixes

Epipolar geometry depends on the relative disposition of two cameras (position and orientation) and internal orientation parameters – principal point coordinates and principal distance (camera constant f_c). The mathematical description of epipolar geometry is based on the usage of fundamental and essential matrixes. Every point from the first image creates a line in other image, on which the corresponding point must lie. This produces constrain which reduces the search in area to search on line. This constrain is due to co-planarity of vectors connecting image points and projection centers and baseline. There are considered two coordinates systems with their origins at the projection centers as it is shown in Fig. 1.

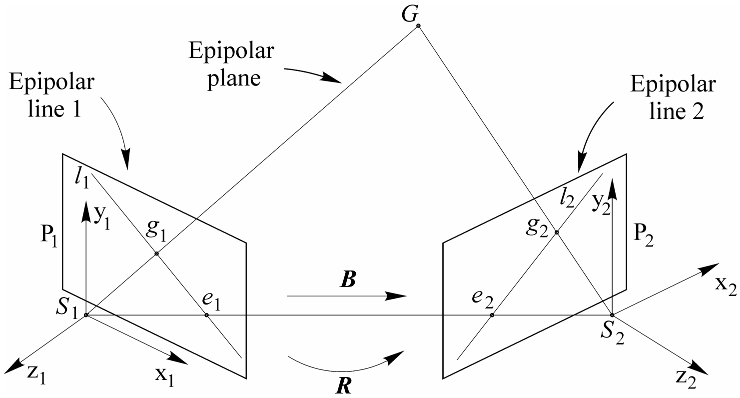

Figure 1. The formulation of fundamental and essential matrixes

Two coordinate systems are connected by rotation matrix $\boldsymbol{R}$ and translation vector $\boldsymbol{B}$.

$$\boldsymbol{g}_2 = \boldsymbol{R}.\boldsymbol{g}_1 + \boldsymbol{B} \ , \tag{3}$$

where $\boldsymbol{g}_1 = \begin{bmatrix} x_1 & y_1 & -f_c \end{bmatrix}^t$ and $\boldsymbol{g}_2 = \begin{bmatrix} x_2 & y_2 & -f_c \end{bmatrix}^t$ are spatial coordinates of object point in two images, $\boldsymbol{B} = \begin{bmatrix} b_x & b_y & b_z \end{bmatrix}^t$ is a base vector, which translate projection center S_1 to position of projection center S_2, and $\boldsymbol{R}$ is a rotation matrix which transforms left image to the orientation of right image.

The relation between images' coordinates and parameters of orientation are expressed by scalar triple product, which expresses the coplanarity condition:

$$\boldsymbol{g}_2^t.(\boldsymbol{B} \times \boldsymbol{R}.\boldsymbol{g}_1) = 0 \ . \tag{4}$$

Re-writing the expression (5) in vector-matrix form leads to expression:

$$\boldsymbol{g}_2^t.\boldsymbol{S}.\boldsymbol{R}.\boldsymbol{g}_1 = 0 \ , \tag{5}$$

where $\boldsymbol{S}$ is base matrix, which is created by the conversion of vector product into matrix form:

$$\boldsymbol{S} = \begin{bmatrix} 0 & -b_z & b_y \\ b_z & 0 & -b_x \\ -b_y & b_x & 0 \end{bmatrix} . \tag{6}$$

Finally the equation (5) in vector-matrix form obtains the presentation:

$$\begin{bmatrix} x_2 & y_2 & z_2 \end{bmatrix}.\boldsymbol{E}.\begin{bmatrix} x_1 & y_1 & z_1 \end{bmatrix}^t = 0 \qquad \boldsymbol{g}_2^t.\boldsymbol{E}.\boldsymbol{g}_1 = 0 \ , \tag{7}$$

$$\text{where } \boldsymbol{E} = \boldsymbol{S}.\boldsymbol{R} \ . \tag{8}$$

Matrix $\boldsymbol{E}$ is named essential matrix, which expresses in mathematical form the epipolar geometry, when parameters of inner orientation are known. Essential matrix is defined by 6 parameters of relative orientation but it is normalized by scale factor and finally has 5 degrees of freedom.

3. Inner camera calibration

3.1. Coordinate system of pinhole camera

Traditional formulation defined by Faugeras [2] uses right coordinate system which x and y axes are oriented to coincide with the direction of increasing the number of columns and rows and z axis is directed toward the shooting direction. This orientation ensures the simple form of projection matrix. But such definition does not correspond to the traditional orientation of coordinate systems in aerial or close-range photogrammetry.

3.2. Definition of orientation for aerial photogrammetry

The calibration of camera is defined for coordinate system of aerial photogrammetry. The origin of coordinate system coincides with projection center of camera S, XY axes are parallel to image coordinate system (x,y), and axis Z is oriented in opposite direction of shooting direction. Point in object space G is projected onto the image plane in point g, in coordinate system with origin the principal point of image [3]. The relations between coordinates in two coordinate system (O,x,y) and (S,X,Y,Z) are given by the equations:

$$x = -\frac{f_c.X}{Z} \qquad y = -\frac{f_c.Y}{Z} \quad . \tag{9}$$

The actual pixel coordinates in left coordinate system (u,v) with origin at the upper left corner (as shown in Fig. 2) are presented by relations:

$$u = u_c + \frac{x + x_0}{\Delta x} \qquad v = v_c - \frac{y + y_0}{\Delta y} \quad , \tag{10}$$

where (u_c, v_c) are pixel coordinates of center of the image, which are defined by the relations: $u_c = P/2 + 0.5$, $v_c = L/2 + 0.5$, where P and L are numbers of pixels and rows, $(\Delta x, \Delta y)$ are raster steps in x and y directions (x_0, y_0) are coordinates of principal point.

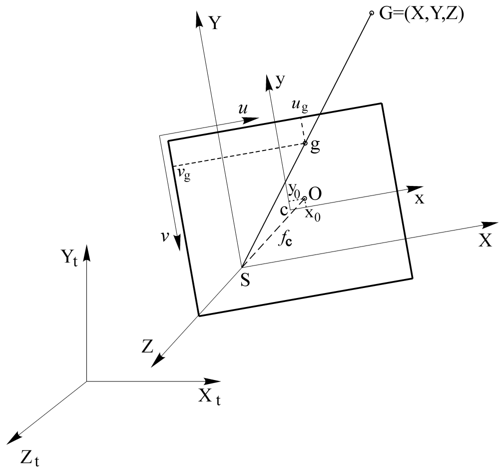

Figure 2. Coordinate system for calibration of aerial camera

The replacement of x and y from (9) in (10) and multiplying with Z leads to the expressions:

$$\begin{aligned} Zu &= Zu_c + \frac{-X.f_c + Z.x_0}{\Delta x} \\ Zv &= Zv_c - \frac{-Y.f_c + Z.y_0}{\Delta y} \end{aligned} . \tag{11}$$

The final presentation in matrix notation has the form:

$$\begin{bmatrix} hu \\ hv \\ h \end{bmatrix} = \begin{bmatrix} k_u & 0 & -u_0 & 0 \\ 0 & -k_v & -v_0 & 0 \\ 0 & 0 & -1 & 0 \end{bmatrix} . \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \qquad \tilde{\boldsymbol{m}} = \tilde{\boldsymbol{P}}^{(a)}.\tilde{\boldsymbol{G}} \ . \tag{12}$$

where $\tilde{\boldsymbol{P}}^{(a)}$ is matrix of perspective transformation in homogeneous form. It could be presented in the following notation: $\tilde{\boldsymbol{P}}^{(a)} = \begin{bmatrix} \boldsymbol{P}^{(a)} & \boldsymbol{q} \end{bmatrix}$, where $\boldsymbol{P}^{(a)}$ is matrix of projective transformation and $\boldsymbol{q}$ is a translation vector.

For scale factor of homogeneous coordinates is obtained the value $h = -Z$.

Parameters (u_c, v_c) and (x_0, y_0) form pixel coordinates of principal point (u_0, v_0) relatively to upper left corner of the image:

$$u_0 = u_c + x_0 / \Delta x \qquad v_0 = v_c - y_0 / \Delta y \ , \tag{13}$$

The five parameters of inner orientation are camera constant f_c, Δx, Δy and pixel coordinates of principal point (u_0, v_0). Only four of them are independent because the arbitrary scale factor could be used for ratio of f_c and pixel size. The scale factors k_u, k_v are defined by relations:

$$k_u = f_c / \Delta x, \quad k_v = f_c / \Delta y \ . \tag{14}$$

Four parameters k_u, k_v, u_0 и v_0 does not depend from position and camera orientation and they represent the parameters of inner orientation.

Camera is a system, which accomplishes linear projective transformation from three-dimensional object space $\mathcal{P}^3$ into two-dimensional projective space $\mathcal{P}^2$.

3.3. Definition of orientation for close-range photogrammetry

The similar formulation can be used for camera system in close-range photogrammetry. The formulation of projection matrix is based on the usage of right coordinate system with x axis which is oriented to the rightwards, z axis oriented upwards and y axis coincide with shooting direction to the object.

$$\begin{aligned} Yu &= Yu_c + \frac{X.f_c + Y.x_0}{\Delta x} \\ Yv &= Yv_c - \frac{Z.f_c + Y.z_0}{\Delta z} \end{aligned} \tag{15}$$

For such definition of coordinate system the equation for projective matrix obtains the form:

$$\begin{bmatrix} hu \\ hv \\ h \end{bmatrix} = \begin{bmatrix} k_u & u_0 & 0 & 0 \\ 0 & v_0 & -k_v & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} . \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \qquad \tilde{\boldsymbol{m}} = \tilde{\boldsymbol{P}}^{(c)} . \tilde{\boldsymbol{G}} \tag{16}$$

Two-dimensional coordinates do not depend on the position and orientation of camera but only of its parameters of inner orientation.

Parameters (u_c, v_c) and (x_0, z_0) form pixel coordinates of principal point (u_0, v_0) relatively to upper left corner of the image:

$$u_0 = u_c + x_0 / \Delta x \qquad v_0 = v_c - z_0 / \Delta z \ , \tag{17}$$

The five parameters of inner orientation are camera constant f_c, Δx, Δz and pixel coordinates of principal point (u_0, v_0). Only four of them are independent because the arbitrary scale factor could be used for ratio of f_c and pixel size. The scale factors k_u, k_v are defined by relations:

$$k_u = f_c / \Delta x, \quad k_v = f_c / \Delta z \ . \tag{18}$$

Four parameters k_u, k_v, u_0 и v_0 does not depend from position and camera orientation and they represent the parameters of inner orientation.

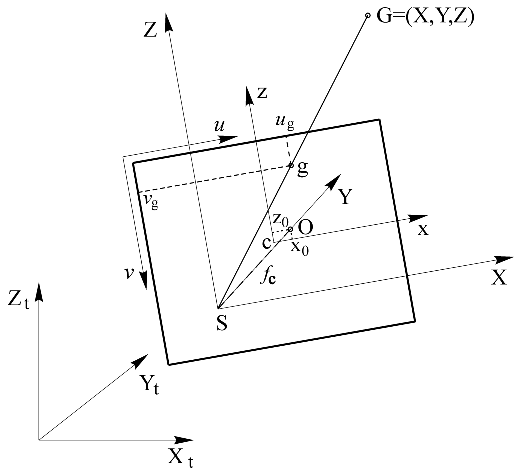

Figure 3. Coordinate system for calibration of close-range camera

4. Relation between fundamental and essential matrixes

Image points and rays in Euclidean 3-D space are presented by the relations:

$$\tilde{\boldsymbol{m}}_1 = \boldsymbol{P}_1.\boldsymbol{g}_1 = \boldsymbol{P}_1 \begin{bmatrix} x_1 \\ y_1 \\ -f_c \end{bmatrix} \qquad \tilde{\boldsymbol{m}}_2 = \boldsymbol{P}_2.\boldsymbol{g}_2 = \boldsymbol{P}_2 \begin{bmatrix} x_2 \\ y_2 \\ -f_c \end{bmatrix} . \qquad (19)$$

Relations (17) allow pixel homogeneous coordinates to be expressed by spatial coordinates and perspective projection matrixes $\boldsymbol{P}_1$ and $\boldsymbol{P}_2$.

$$\boldsymbol{g}_1 = \boldsymbol{P}_1^{-1}\tilde{\boldsymbol{m}}_1 \qquad \boldsymbol{g}_2 = \boldsymbol{P}_2^{-1}\tilde{\boldsymbol{m}}_2 \qquad (20)$$

The replacement of expressions for $\boldsymbol{g}_1$ and $\boldsymbol{g}_2$ in equation (7) leads to the form of the equation where the expression for fundamental matrix, image point $\tilde{\boldsymbol{m}}_1$ and its corresponding point $\tilde{\boldsymbol{m}}_2$ are connected by the relation:

$$\tilde{\boldsymbol{m}}_2^t.\left(\boldsymbol{P}_2^{-1}\right)^t.\boldsymbol{E}.\boldsymbol{P}_1^{-1}.\tilde{\boldsymbol{m}}_{\mathbf{1}} = 0 \ . \qquad (21)$$

Finally the fundamental matrix obtains the presentation which depends on essential matrix and perspective projective matrixes of inner orientation of cameras:

$$\boldsymbol{F} = \left(\boldsymbol{P}_2^{-1}\right)^t.\boldsymbol{E}.\boldsymbol{P}_1^{-1} \ . \qquad (22)$$

The fundamental matrix has dimension 3×3 but because it is defined with arbitrary scale factor and satisfies the constraint $|\boldsymbol{F}| = 0$ that leaves for it only 7 degrees of freedom and it has rank 2.

5. Discussions and conclusions

The usage of epipolar geometry allows mathematical formulation of relations between corresponding epipolar lines in two stereo images without usage of spatial coordinates of control points. The inner parameters of camera orientation could be determined by the measurements of image points on the same object which are registered in images taken from three or more different camera positions. The epipolar geometry is important for finding the correspondence of image points for measurements of spatial coordinates and for automatic forming of object model in digital photogrammetry.

REFERENCES

1. *Faugeras, Olivier*. What can be seen in three dimensions with an uncalibrated stereo rig? Proc. *ECCV 92,* Lecture Notes in Computer Science, 588, Springer-Verlag, 1992.
2. *Faugeras, Olivier*. Three-Dimensional Computer Vision: A Geometric Viewpoint. M.I.T. Press, 1993.
3. *Marinov, B*. Cybernetics and Information Approaches in Photogrammetry. Sofia, Classic Design, 2012 (in Bulgarian).
4. *Maybank Stephen, J. and Olivier D. Faugeras*. A theory of self-calibration of a moving camera. International Journal of Computer Vision, 8:2, 123-151, 1992.