EPIPOLAR GEOMETRY BETWEEN PHOTOGRAMMETRY AND COMPUTER VISION – A COMPUTATIONAL GUIDE

Bashar Alsadik [a*], Nagham Amer Abdulateef [b]

[a*] Faculty of Geo-Information Science and Earth Observation (ITC), University of Twente, Enschede, The Netherlands - b.s.a.alsadik@utwente.nl

[b] Dept. of surveying, College of Engineering, University of Baghdad, Iraq - n.a.mohammed@coeng.uobaghdad.edu.iq

Commission V, WG V/7

KEY WORDS: coplanarity, collinearity, fundamental matrix, essential matrix, relative orientation, Photogrammetry.

ABSTRACT:

Stereo image orientation is one of the major topics in computer vision, photogrammetry, and robotics. The stereo vision problem solution represents the basic element of the multi-view Structure from Motion SfM in computer vision and photogrammetry.
A successfully reconstructed stereo image geometry is based on solving the epipolar constraint using the fundamental matrix which is based on the projective geometry in computer vision. However, in photogrammetry, the problem is well known as relative orientation and there is a different solution that is based on the euclidean geometry using collinearity or coplanarity equations.
A lot of literature and discussions were found in the last decades to solve the epipolar geometry problem. However, there is still no clear description to compare between solutions introduced using both projective and euclidean solutions and which method of the relative image orientation is mostly preferred.
To the best of our knowledge, computing and plotting the epipolar lines using photogrammetric collinearity and coplanarity equations is not shown before in the educational litrature. In this paper, a detailed mathematical solution of the epipolar geometry will be shown using both photogrammetric and computer vision techniques. This is aimed to remove any confusion for new learners in using the current methods in both scientific fields and show that using any technique should lead to comparable results with advantages and disadvantages.

1. INTRODUCTION

A long discussion has happened in the last few decades between both communities of photogrammetry and computer vision since they both aim for solving similar vision problems in different ways. In photogrammetry, the aim was started for topographic mapping and military reconnaissance and that pushed to reach a high level of precision based on using calibrated cameras (Forstner, 2002). While with the advanced development of computer science, computer or machine vision gained a lot of attention and was enriched with a wide range of techniques that handles the same problems that were solved by photogrammetry. Interestingly, photogrammetry mathematical techniques were based on Euclidean geometry while computer vision used homogeneous coordinates and projective geometry. Generally, more direct solutions were found in the computer vision field assuming mostly uncalibrated cameras while assuming calibrated cameras with the development of nonlinear solutions in photogrammetry. This use of direct solutions in computer vision using the algebraic minimization didn't mostly imply the uncertainty in the measurements (Forstner, 2002). On the other hand, the photogrammetric nonlinear rigorous solutions are very much relying on the starting values of the unknowns which is a nontrivial task. Remarkably, most photogrammetric packages nowadays are either using computer vision techniques to handle the uncalibrated camera cases or a mixture of both to count also for apriori uncertainty of measurements and estimate final precision.
For students and specialists from both fields, it's not an easy task to understand the techniques adopted by both fields because of the different types of mathematical treatment behind them. Accordingly, in this paper, one aim is to introduce a computational guide to solve the important problem of the epipolar geometry of a stereo pair of images using the Euclidean and projective geometrical solutions and assess their performances with the existence of noise or blunders.
The euclidean-based nonlinear solution adopted for a long time in photogrammetry is the well-known *collinearity equations* (Liu et al., 2006; Luhmann T. et al., 2006; Salma, 1980; Wolf and DeWitt, 2000) which is the workhorse for solving the image orientation problems. Another condition equation called the *coplanarity* is also developed for a long time in photogrammetry mainly to recover the stereo image orientation problem using euclidean coordinates (Salma, 1980; Wolf and DeWitt, 2000). From a computer vision perspective, the collinearity equations are defined as the pinhole camera model using homogenous coordinates (Forstner, 2002; Hartley and Zisserman, 2003) where a special matrix called the projection matrix P is developed. Worth mentioning that the projection matrix was recognized by the photogrammetry field for five decades under the name of Direct Linear Transformation DLT (Luhmann, 2014; Salma, 1980).
When the vision problem extends from the single image geometry to stereo, epipolar constraint arise and constitute the base of the solution. The epipolar geometry entails that for a stereo camera configuration, the center of each camera $O1$, $O2$, and the world point P defines a plane in space which is called the epipolar plane (Figure 1a). The result of projecting the world point P onto the image planes of the two cameras is to have two conjugate points $p1$ and $p2$ respectively. Furthermore, two special points e_1 and e_2 called the epipoles will be found as a result of the projection of each camera into the other. The intersection of the epipolar plane with the two image planes will result in two epipolar lines $l1$ and $l2$ where they pass through e_1

https://doi.org/10.5194/isprs-annals-V-5-2022-25-2022

and e_2 respectively. Every point in one image will have its conjugate in the second image constrained to lie along the epipolar line (Figure 1b). The epipolar constraint has a great benefit in dens image matching, image rectification, and Structure from Motion SfM among other applications.
In literature, it's difficult to find a detailed description for solving the mentioned epipolar constraint and the relative orientation of the same problem using both euclidean and homogenous coordinates. Some papers were presented to investigate which method can provide more accurate solutions. For example, the authors (Kim and Kim, 2016) presented the solution of the epipolar geometry applied in computer vision and photogrammetry and concluded that despite the epipolar resampling methods developed in the two fields being mathematically identical, their performance in the epipolar parameter estimation may be changed.

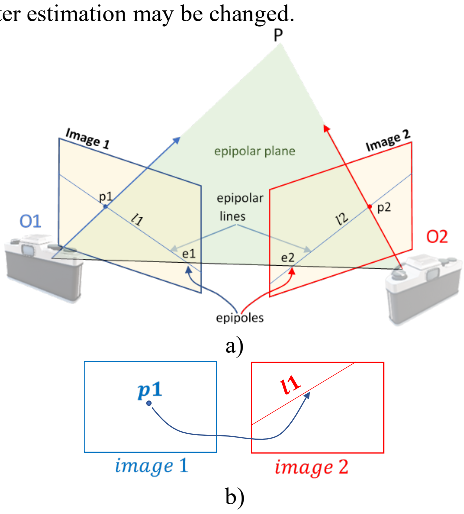

Figure 1. Epipolar geometry.

In (El-Ashmawy, 2015) it was shown that using the coplanarity and collinearity will give more accurate results of the estimated camera exterior orientation compared to the DLT method which is based on the projective geometry. This conclusion is assured by (Barrile et al., 2017) who found that using photogrammetric relative and absolute orientation of a triple of images can give more accurate results than the computer vision conventional techniques.
This paper aims to give a detailed solution for the stereo vision epipolar constraint given in computer vision and photogrammetry using both homogeneous and euclidean coordinates. This is expected to help the learners from both fields to get familiar with the solution introduced by the other field. In the end, it is hoped that an integrated solution that is more robust and less error-prone can be found.
The paper structure starts after this introduction with a brief presentation of the method in section 2 and then continues in section 3 to present the detailed mathematical solutions using the euclidean and homogenous coordinates. Then, section 4 illustrates the results and completes them with discussions. Conclusions will be finally given in section 5.

2. METHOD

Three techniques using coplanarity, collinearity, and fundamental matrix will be investigated to achieve the relative orientation task for the same stereo images data in two different experiments. The numerical solutions will be given for the benefit of reproducing the calculations and applying a more sophisticated analysis by the readers.

To understand the mathematical concepts behind the mentioned techniques, we will describe them with illustrations in the following section 2.1 explaining the epipolar geometry processing from both perspectives of traditional photogrammetry and section 2.2 describing the computer vision approach using the fundamental matrix.

2.1 Epipolar constraint using euclidean geometry

In conventional photogrammetry, mathematical computations are applied using the euclidean coordinates. So we have the image coordinates defined as $[x, y]$, and the object coordinates as $[X, Y, Z]$ coordinates. Two mathematical techniques are found to solve the stereo relative orientation, the collinearity, and the coplanarity conditions. The limitation of these two techniques is to have calibrated camera parameters known.

A. Coplanarity Equations

Coplanarity equation 1 is based on the condition of having any object point P, the left camera $O1$, the right camera $O2$, and the corresponding image points $p1$, $p2$ on the two images to compose one plane (Schindler, 2015) as shown in Figure 2. Assuming a rotation matrix R and translation T for every camera are given, then the coplanarity condition can be formulated as a dependent relative orientation problem where the first image is assumed fixed while orienting the second image concerning it. Accordingly, a plane equation in 3D can be formulated as a determinant G of a 3x3 matrix as follows:

$$G = det\begin{bmatrix} bx & by & bz \\ x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \end{bmatrix} \qquad (1)$$

where

$$T = T^2 - T^1 = [bx \quad by \quad bz]^t, R^1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, R^2 = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \qquad (2)$$

The two vectors pointing from the two cameras to the object point can be written in terms of the scale factors s (assumed equal in the two images), rotation matrices R, and the image coordinates p as $s_1 {R^1}^t \vec{p}_2$ and $s_2 {R^2}^t \vec{p}_2$. Therefore

$$\begin{bmatrix} x_1 \\ y_1 \\ z_1 \end{bmatrix} = {R^1}^t \begin{bmatrix} x_{p1} \\ y_{p1} \\ -f \end{bmatrix}, \quad \begin{bmatrix} x_2 \\ y_2 \\ z_2 \end{bmatrix} = {R^2}^t \begin{bmatrix} x_{p2} \\ y_{p2} \\ -f \end{bmatrix} \qquad (3)$$

In the case of the unknown orientation of the two stereo images, the coplanarity equations should be applied to solve the relative rotation R^2 and translation T^2 of the second image while assuming the first image fixed $R^1 = I_{3*3}, T^1 = [0,0,0]^t$. This is conventionally known in photogrammetry as the dependent relative orientation problem.

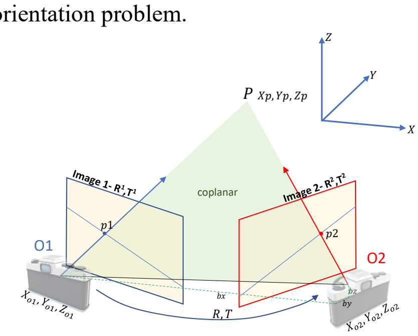

Figure 2. Coplanarity condition.

The solution starts by either manually measuring corresponding points distributed well on the overlap area or estimating the matching points by using one of the well-known operators like SIFT. To proceed, the pixel coordinates of the matching points should be transformed to the principal point $p.p.$ coordinate-system (image center) following equation 6.
Then an iterative nonlinear least-squares solution should be applied. If **Euler angles** ω_2, φ_2, k_2 are used to define the rotation matrix R^2, then the starting values are normally set up as zero while a base component bx is fixed to 1 for example in the translation vector $T^2 = [1 \quad by \quad bz]^t$ to fix the scale of the problem. Obviously, a minimum of five corresponding points is required to solve the coplanarity condition (Wolf and DeWitt, 2000) where the observation equations can be set up in a matrix form as $AV + B\Delta = L$ or:

$$\underbrace{\begin{bmatrix} \frac{\partial G_1}{\partial x_{p1}{}^1} & \frac{\partial G_1}{\partial y_{p1}{}^1} & \frac{\partial G_1}{\partial x_{p2}{}^1} & \frac{\partial G_1}{\partial y_{p2}{}^1} & & & & & & & & 0 \\ & & & & \frac{\partial G_2}{\partial x_{p1}{}^2} & \frac{\partial G_2}{\partial y_{p1}{}^2} & \frac{\partial G_2}{\partial x_{p2}{}^2} & \frac{\partial G_2}{\partial y_{p2}{}^2} & & & & \vdots \\ & & 0 & & & \cdots & & & \frac{\partial G_n}{\partial x_{p1}{}^n} & \frac{\partial G_n}{\partial y_{p1}{}^n} & \frac{\partial G_n}{\partial x_{p2}{}^n} & \frac{\partial G_n}{\partial y_{p2}{}^n} \end{bmatrix}}_{A_{n\times 4n}} \underbrace{\begin{bmatrix} v_{x_{p1}}{}^1 \\ v_{y_{p1}}{}^1 \\ v_{x_{p2}}{}^1 \\ v_{y_{p2}}{}^1 \\ \vdots \\ v_{y_{p2}}{}^n \end{bmatrix}}_{V_{4n\times 1}} +$$

$$\underbrace{\begin{bmatrix} \frac{\partial G_1}{\partial \omega_2} & \frac{\partial G_1}{\partial \varphi_2} & \frac{\partial G_1}{\partial k_2} & \frac{\partial G_1}{\partial by} & \frac{\partial G_1}{\partial bz} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{\partial G_n}{\partial \omega_2} & \frac{\partial G_n}{\partial \varphi_2} & \frac{\partial G_n}{\partial k_2} & \frac{\partial G_n}{\partial by} & \frac{\partial G_n}{\partial bz} \end{bmatrix}}_{B_{n\times 5}} \underbrace{\begin{bmatrix} \delta_{\omega_2} \\ \delta_{\varphi_2} \\ \delta_{k_2} \\ \delta b_y \\ \delta b_z \end{bmatrix}}_{\Delta_{5\times 1}} = \underbrace{\begin{bmatrix} G_1 \\ G_2 \\ \vdots \\ G_n \end{bmatrix}}_{L_{n\times 1}} \qquad (4)$$

where the B matrix refers to the partial derivatives of G (equation 1) concerning the unknowns, A matrix refers to the partial derivatives of G concerning the image points observations, V is the residuals vector, Δ is the required corrections vector, and L is the functional values vector.
Then to handle the redundancy and nonlinearity, a least-squares adjustment is applied using equation 5 to estimate the corrections Δ as follows (Alsadik, 2019):

$$\Delta = (B^t(AQA^t)^{-1}B)^{-1}B^t(AQA^t)^{-1}L \qquad (5)$$

Where Q refers to the cofactor weight matrix which is mostly selected as an identity matrix assuming all the image coordinates are of equal weight since they are detected by the same operator like SIFT. Accordingly, the corrections will be updated iteratively until reaches an insignificant number (like $\leq 10^{-7}$).

Then to estimate and plot the epipolar constraint between the stereo images, it is possible to apply the following steps:

- transform the pixel coordinates to the $p.p.$ system of the first image assuming the coordinates of the $p.p.$ are x_o, y_o

$$x_{p1} = x_1 - \frac{image\ width}{2} - 0.5 - x_o$$
$$y_{p1} = \frac{image\ height}{2} - 0.5 - y_1 - y_o \qquad (6)$$

Worth noting, format width, height, and focal length can be used in pixels or mm units.

- apply the estimation of the following parameters based on the measured image point, the base components bx, by, bz, and rotation elements $r's$.

$$c = -bz\, r_{32}\, x1\, f + bz\, r_{31}\, y1\, f - r_{31}\, z1\, by\, f + r_{33}\, f\, x1\, by - r_{33}\, f\, y1\, bx + r_{32}\, f\, z1\, bx \qquad (7)$$
$$a = r_{12}\, x1\, bz - r_{11}\, y1\, bz + r_{11}\, z1\, by - r_{13}\, x1\, by + r_{13}\, y1\, bx - r_{12}\, z1\, bx \qquad (8)$$
$$b = r_{22}\, x1\, bz - r_{21}\, y1\, bz + r_{21}\, z1\, by - r_{23}\, x1\, by + r_{23}\, y1\, bx - r_{22}\, z1\, bx \qquad (9)$$

- Randomly select some corresponding points like $p2$ on the 2nd image in the range [-image width:+image width].
  - Then estimate the e' and n' coordinates of the epipolar line as follows:

$$n' = \frac{(a\, x_{p2} + c)}{-b}, e' = x_{p2} + width/2 \qquad (10)$$

Then trim the points if they are located outside the image format size.

B. Collinearity Equations

The collinearity equations are developed based on the concept of having an object point $[Xp, Yp, Zp]$, camera lens $[X_o, Y_o, Z_o]$, and the corresponding image point $[x_p, y_p]$ located at the same line in space (Luhmann T. et al., 2006; Salma, 1980; Schindler, 2015; Wolf and DeWitt, 2000) Figure 3.

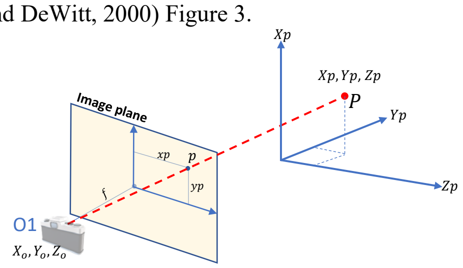

Figure 3. Collinearity condition.

Collinearity equations are algebraically formulated as follows using Euclidean geometry:

$$\left.\begin{aligned} G_{x_p} &= -f\frac{r_{11}(X_p-X_o)+r_{12}(Y_p-Y_o)+r_{13}(Z_p-Z_o)}{r_{31}(X_p-X_o)+r_{32}(Y_p-Y_o)+r_{33}(Z_p-Z_o)} - x_p \\ G_{y_p} &= -f\frac{r_{21}(X_p-X_o)+r_{22}(Y_p-Y_o)+r_{23}(Z_p-Z_o)}{r_{31}(X_p-X_o)+r_{32}(Y_p-Y_o)+r_{33}(Z_p-Z_o)} - y_p \end{aligned}\right] \qquad (11)$$

where X_o, Y_o, Z_o= camera coordinates.
f: focal length in mm or pixels.
x_p, y_p: image coordinates in mm or pixels.
$r's$=rotation matrix elements as $R = R_k R_\varphi R_\omega$ or:

$$R = \begin{bmatrix} \cos\varphi\cos k & \cos\omega\sin k + \sin\omega\sin\varphi\cos k & \sin\omega\sin k - \cos\omega\sin\varphi\cos k \\ -\cos\varphi\sin k & \cos\omega\cos k - \sin\omega\sin\varphi\sin k & \sin\omega\cos k + \cos\omega\sin\varphi\sin k \\ \sin\varphi & -\sin\omega\cos\varphi & \cos\omega\cos\varphi \end{bmatrix} \qquad (12)$$

It should be noted that the rotations are based on a right-handed system.
Similarly to the coplanarity, a minimum of five corresponding points is required to solve the relative orientation using collinearity. So if points a, b, c, d, and g are used in solving the relative orientation problem then every point will involve two collinearity equations and then the observation equations (Wolf and DeWitt, 2000) will be represented in matrix form as shown in equation 13 where the gray refers to image 1 and yellow refers to the image 2:

$$\begin{bmatrix} B_a^s & & & & 0 \\ & B_b^s & & & \\ & & B_c^s & & \\ & & & B_d^s & \\ 0 & & & & B_g^s \\ B_a^e & B_a^s & & & 0 \\ B_b^e & & B_b^s & & \\ B_c^e & & & B_c^s & \\ B_d^e & & & & B_d^s \\ B_g^e & 0 & & & B_g^s \end{bmatrix} \begin{bmatrix} \overset{e}{\Delta} \\ \overset{s}{\Delta_a} \\ \overset{s}{\Delta_b} \\ \overset{s}{\Delta_c} \\ \overset{s}{\Delta_d} \\ \overset{s}{\Delta_g} \end{bmatrix} - \begin{bmatrix} v_{xa} \\ v_{ya} \\ \vdots \\ v_{yg} \\ v_{xa} \\ v_{ya} \\ \vdots \\ v_{yg} \end{bmatrix} = \begin{bmatrix} G_{xa} \\ G_{ya} \\ \vdots \\ G_{yg} \\ G_{xa} \\ G_{ya} \\ \vdots \\ G_{yg} \end{bmatrix} \qquad (13)$$

where e refers to exterior orientation parameters.
s refers to the object points defined in $[X, Y, Z]$.
Δ refers to the corrections, then
$\overset{e}{\Delta} = [\delta\omega_2\ \delta\varphi_2\ \delta k_2\ \delta by\ \delta bz]^t$, and for point a the correction vector is $\overset{s}{\Delta}_a = [\delta X_a\ \delta Y_a\ \delta Z_a]^t$ and so on.

The B matrices representing the jacobian matrices, for example, B_a^e and B_a^s of point a are composed by taking the partial derivation of the collinearity equations 11 to the orientation parameters and the object point a respectively as:

$$B_a^e = \begin{bmatrix} \left(\frac{\partial G_{xa}}{\partial \omega_2}\right) & \left(\frac{\partial G_{xa}}{\partial \varphi_2}\right) & \left(\frac{\partial G_{xa}}{\partial k_2}\right) & \left(\frac{\partial G_{xa}}{\partial by}\right) & \left(\frac{\partial G_{xa}}{\partial bz}\right) \\ \left(\frac{\partial G_{ya}}{\partial \omega_2}\right) & \left(\frac{\partial G_{ya}}{\partial \varphi_2}\right) & \left(\frac{\partial G_{ya}}{\partial k_2}\right) & \left(\frac{\partial G_{ya}}{\partial by}\right) & \left(\frac{\partial G_{ya}}{\partial bz}\right) \end{bmatrix}$$

$$B_a^s = \begin{bmatrix} \left(\frac{\partial G_{xa}}{\partial X_a}\right) & \left(\frac{\partial G_{xa}}{\partial Y_a}\right) & \left(\frac{\partial G_{xa}}{\partial Z_a}\right) \\ \left(\frac{\partial G_{ya}}{\partial X_a}\right) & \left(\frac{\partial G_{ya}}{\partial Y_a}\right) & \left(\frac{\partial G_{ya}}{\partial Z_a}\right) \end{bmatrix} \tag{14}$$

Then the least-squares adjustment (equation 5) is applied to compute the most probable values of the second image relative orientation besides the object points coordinates.
When the orientation of the two images is solved, epipolar lines constraint can be constructed. The mathematical approach to solve the epipolar constraint using collinearity can be described as follows:

- transform the pixel coordinates to the principal point $p.p.$ coordinates of the first image (equation 6).
- Apply the inverse collinearity equations of the form

$$X = X_o + (Z - Z_o)\frac{m_{11}(x - x_o) + m_{21}(y - y_o) + m_{31}(-f)}{m_{13}(x - x_o) + m_{23}(y - y_o) + m_{33}(-f)}$$
$$Y = Y_o + (Z - Z_o)\frac{m_{12}(x - x_o) + m_{22}(y - y_o) + m_{32}(-f)}{m_{13}(x - x_o) + m_{23}(y - y_o) + m_{33}(-f)} \tag{15}$$

where Z value will be ranged arbitrarily like between 10 and -10.

- After computing the ground XY coordinates associated with the assumed values of Z on the first image, The same coordinates will be projected to the second image which has its orientation parameters using the forward collinearity equations (11).
- Finally, transform the coordinates from the principal point system to the pixel coordinate system.
- The x, y image coordinates will be fitted to a 2D line which will represent the epipolar line on the second image. This approach can be inverted for finding the epipolar lines associated with the image points at the 2nd image.

2.2 Epipolar constraint using projective geometry

In computer vision, mathematical computations are applied using homogenous coordinates. So we have the image coordinates defined as $[x, y, 1]^t$, and the object coordinates as $[X, Y, Z, 1]^t$ coordinates. The *epipolar geometry* of two stereo images is described by a very special singular 3x3 matrix called the *fundamental matrix* F. The F matrix expresses the relative orientation between the two stereo frames similarly to coplanarity condition equations. Worth mentioning, F is used to map corresponding conjugate points in image 1 to image 2 where the epipolar line $l2$ of point $p1$ is $Fp1$ and that epipolar line $l1$ of point $p2$ is $F^t p2$ (Hartley and Zisserman, 2003; Szeliski, 2011). Therefore F explains the epipolar constraint between two conjugate points (Figure 4) as $p2^t F p1 = 0$ or:

$$p2^t F p1 = [x'_1 \quad y'_1 \quad 1] \begin{bmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} = 0 \tag{16}$$

The F matrix can be computed either by knowing the intrinsic K and extrinsic parameters R, T between the two cameras or when enough corresponding points between the two images are known (uncalibrated camera).
In the uncalibrated camera case, the conventional method is applied by using the 8-points algorithm (equation 17) which is a linear solution that enforces the matrix to have a rank of 2 (Figure 4). This algorithm solution can have difficulties, especially with noisy points and therefore RANSAC technique is mostly used (Szeliski, 2011). Interestingly, using the F matrix without knowing the intrinsic parameters K leads to an uncalibrated SfM solution.

$$AF = \begin{bmatrix} x'_1 x_1 & x'_1 y_1 & x'_1 & y'_1 x_1 & y'_1 y_1 & y'_1 & x_1 & y_1 & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x'_n x_n & x'_n y_n & x'_n & y'_n x_n & y'_n y_n & y'_n & x_n & y_n & 1 \end{bmatrix} \begin{bmatrix} F_{11} \\ F_{12} \\ F_{13} \\ F_{21} \\ F_{22} \\ F_{23} \\ F_{31} \\ F_{32} \\ F_{33} \end{bmatrix} = 0 \tag{17}$$

where n is the number of corresponding points. The system of equations (17) is solved using linear least squares by the singular value decomposition SVD (Hartley and Zisserman, 2003) and selecting the last eigenvector of V.
Preferably, the point coordinates are normalized to avoid the ill-conditioned matrix by firstly locating the origin of the new coordinates at the centroid of the image points (translation) and secondly, by transforming the average distance of the transformed image points from the origin to be like $\sqrt{2}$ pixels (scaling) (Hartley and Zisserman, 2003). The SVD method will result in an estimate of the fundamental matrix, which may be of rank 3 and then an optimization approach should be followed to find the correct fundamental matrix of rank 2.
In the case of having a calibrated camera, equation 18 is showing the estimation method of F as follows:

$$F = K'^{-T}[t]_x R K^{-1} \tag{18}$$

Where

$[t]x = \begin{bmatrix} 0 & -bz & by \\ bz & 0 & -bx \\ -by & bx & 0 \end{bmatrix}$, K is the intrinsic matrix assuming one camera is used.

It worth mentioning that when the intrinsic parameters K are known, then we can compute a singular matrix called the $Essential\ matrix\ E$ which is completely defined by three rotational angles and two translational parameters (dependant relative orientation) as follows (Hongdong and Hartley, 2006; Nister, 2004):

$$E = K^t\, FK = [t]_x R \tag{19}$$

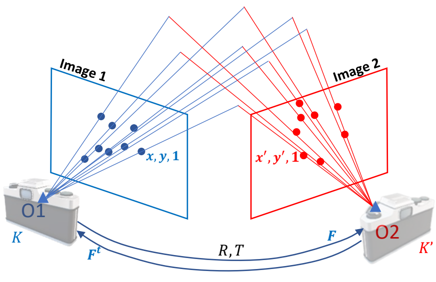

Figure 4. Estimating F matrix using the 8-point algorithm.

The essential matrix can help to avoid the projectivity distortion effect in the uncalibrated SfM and to have a metrical solution.

Conventionally multiple solutions are found and the selection of the most proper matrix should be applied (Forstner, 2002; Hartley and Zisserman, 2003; Szeliski, 2011).
Worth noting, the possibility to decompose E into four possible solutions of the rotation $R_{1,2}$ and translations $t_{1,2}$ using the following algorithm:

- `Decomposition using SVD as [U,D,V] = svd (E)`
- `Prepare` $W1 = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $W2 = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$
- `Translation vector (skew-symmetry matrix) S = UW2U';`
- `Two possible rotation matrices R1 = UWV' and R2 = UW'V'`
- `Two possible translation vectors: t1 = third column of U; t2 = -t1`
- `Check determinant (det):if det(R1)or det(R2)<0, then multiply by -1`

Furthermore, an assessment of the epipolar geometrical reliability will be investigated in the paper using the three presented methods in section 2.1 and section 2.2. Assessment will be applied by computing the distances between the points in one image to the corresponding epipolar lines will be computed using equation 20 (Weisstein, 1996) of the point-line distance as shown in Figure 5.

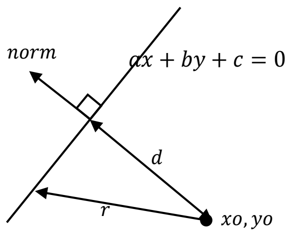

Figure 5. Point-Line Distance [18]

$$d = \frac{|ax_o + bx_o + c|}{\sqrt{a^2 + b^2}} \quad (20)$$

where a, b, and c are the epipolar line equation parameters and xo, yo are the coordinates of the image points while $|\quad|$ refers to the norm.

3. RESULT

To investigate the three computational algorithms, two experimental tests are applied for ground close-range stereo images and aerial images.

3.1 First test

In the first test, two stereo images in an urban environment (Figure 6) are taken using a camera having the following parameters:

- focal length =18 mm
- pixel size= 4.7 µm
- format size= 4753×3168 pixels
- lens distortion is negligible.

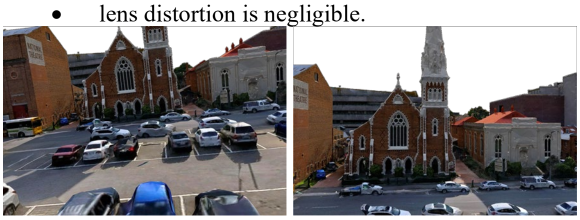

Figure 6. First test stereo images with a significant rotation difference.

A- Using coplanarity equations

The first solution of the relative orientation problem and the epipolar geometry will be investigated using the coplanarity equations given 15 homologous image points (Table 1). Accordingly, the solution is applied and iterated until converged to insignificant correction values. In Table 1 the details of the first iteration least-square matrices are shown.

left image x	left image y	right image x	right image y	1st iteration		B matrix			L matrix		corrections
1527.20	1295.80	1108.80	2196.60	1.6024E+06	8.8518E+05	1.4491E+07	3.6521E+05	-4.8531E+06	3.44990E+06	by	-1.0887
2479.40	956.26	2223.00	1955.30	9.8196E+05	5.7652E+04	1.4434E+07	9.6044E+04	-5.8596E+05	3.82610E+06	bz	-2.1635
1527.30	988.41	1158.90	1860.60	1.4109E+06	9.5964E+05	1.4503E+07	7.2489E+05	-4.6612E+06	3.34030E+06	omega	0.32104
2855.50	1506.20	2651.50	2690.90	7.8128E+05	-5.5219E+05	1.4581E+07	-2.1434E+04	1.0551E+06	4.53710E+06	phi	-0.15211
2802.30	1567.50	2589.40	2754.00	8.1536E+05	-5.0229E+05	1.4648E+07	-3.5211E+03	8.1728E+05	4.54400E+06	kappa	-0.52384
1393.70	1196.40	974.11	2079.40	1.6069E+06	1.0300E+06	1.4475E+07	5.4337E+05	-5.3689E+06	3.38170E+06		
1862.70	1341.40	1488.90	2285.90	1.4316E+06	5.7550E+05	1.4497E+07	2.1521E+05	-3.3974E+06	3.61720E+06		
3742.50	1162.80	3486.90	2314.70	9.7889E+05	-1.4664E+06	1.4359E+07	-4.6791E+05	4.2545E+06	4.41150E+06		intial values
730.52	1630.60	93.57	2538.90	2.4394E+06	1.4649E+06	1.4712E+07	-1.0636E+05	-8.7412E+06	3.47860E+06	by	1
2204.20	754.52	1942.80	1698.70	1.0011E+06	3.7904E+05	1.4572E+07	3.5933E+05	-1.6591E+06	3.61600E+06	bz	0
2909.30	1006.20	2651.40	2041.70	9.8770E+05	-4.0322E+05	1.4403E+07	-1.5913E+05	1.0547E+06	3.96570E+06	omega	0
2442.10	125.22	2265.60	1102.90	6.7596E+05	1.9285E+05	1.5369E+07	1.6105E+05	-4.2281E+05	3.74430E+06	phi	0
2658.80	829.30	2413.70	1833.50	9.3868E+05	-9.9011E+04	1.4479E+07	-2.8452E+04	1.4438E+05	3.84590E+06	kappa	0
1893.30	882.15	1588.90	1794.20	1.1658E+06	6.5389E+05	1.4520E+07	5.5243E+05	-3.0144E+06	3.49300E+06		
2824.90	619.36	2603.00	1641.30	8.4983E+05	-2.4470E+05	1.4612E+07	-2.1897E+05	8.6936E+05	3.91380E+06		

Table 1. The image correspondences and the least-squares adjustment of the first iteration

The nonlinear least-squares adjustment continues and stops when the corrections are negligible (<10^{-5}) as shown in Figure 7.

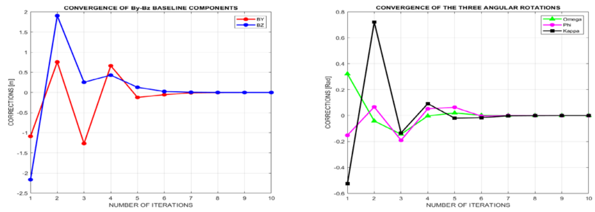

Figure 7. The convergence of the relative orientation parameters corrections using coplanarity – test 1.

Table 2 summarizes the final relative orientation parameters and their precision σ using coplanarity equations.

	ω_{deg}	φ_{deg}	k_{deg}	bY	bZ
Adjusted	8.7923	-9.5087	6.5114	-1.1236	0.5837
σ	±0.0482	±0.0289	±0.0395	±0.0041	±0.0032
Initial	0	0	0	1	0

Table 2. Relative orientation parameters using coplanarity.

Furthermore, the epipolar lines of the same corresponding points of Table 1 are calculated using equations 6-10 as shown in Figure 8.

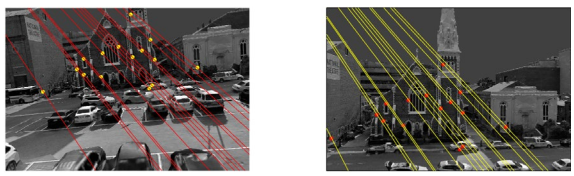

Figure 8. The image point correspondences and the epipolar lines on the stereo images using photogrammetric coplanarity or collinearity equations.

B- Using collinearity equations

The second solution of the relative orientation problem and the epipolar geometry will be investigated using the collinearity equations given the same homologous image points (Table 1). Accordingly, the solution is applied and iterated until converged to insignificant correction values. In Figure 9, the sparse coefficient matrix B and the normal equations matrix of equation 5 are shown. It's worth mentioning that the zero values assumed for the left image orientation have their effect on having zero partial derivatives shown in the matrix of coefficients.

https://doi.org/10.5194/isprs-annals-V-5-2022-25-2022

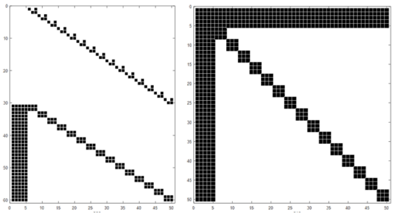

Figure 9. Sparse matrices pattern of the least-squares adjustment using collinearity equations.

The nonlinear least-squares adjustment continues and stops when all the corrections are negligible ($<10^{-5}$) as shown in Figure 10. More than the image orientation computations, collinearity will produce the XYZ coordinates of the object points in an arbitrary scale and at the reference of the first image. These points in a multiview problem can finally result in a sparse point cloud similar to the conventional approach of SfM. However, the proper estimate of their initial coordinates is crucial to have a convergent nonlinear collinearity solution. Table 3 summarizes the final relative orientation parameters using collinearity equations and their precision σ. As logically expected, the same parameter values are estimated using the coplanarity equations.

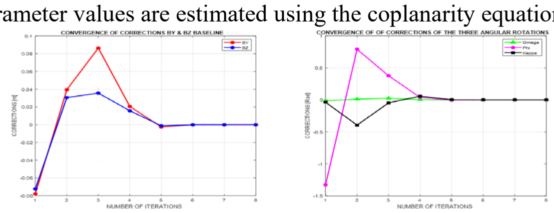

Figure 10. The convergence of the relative orientation parameters corrections using collinearity – test 1.

	ω_{deg}	φ_{deg}	k_{deg}	bY	bZ
Adjusted	8.7924	-9.5092	6.5115	-1.1235	0.5837
σ	±0.0502	±0.0297	±0.0134	±0.0032	±0.0037
Initial	-10	5	0	-1	0

Table 3. Relative orientation parameters using collinearity

Additionally, the epipolar lines of the same corresponding points used to run coplanarity are calculated using equations 11-15 as shown previously in Figure 8.

C- Using the fundamental matrix

The first step here is to consider the homogenous coordinates of the points by adding 1's to the image pixel coordinates.
Then find the transformation that normalizes the image points which can be achieved by

1. Find the mean of the interest points
2. Find the distance of the image points from their mean
3. Normalize all the image points using the mean and distance
4. Find the transformation for normalizing points

Accordingly, two matrices T1 and T2 will be found:

$$T1 = \begin{bmatrix} 0.00127 & & -1.1957 \\ & 0.00127 & -1.4552 \\ & & 1 \end{bmatrix}, T2 = \begin{bmatrix} 0.00111 & & -1.2192 \\ & 0.00111 & -1.5411 \\ & & 1 \end{bmatrix}$$

Then, using Linear Least Squares, solve for an initial estimate of the Fundamental Matrix by determining the SVD of A matrix of equation 17 and selecting the last eigenvector. Denormalizing and conditioning the Fundamental Matrix to enforce the rank constraint (Sarkar, 2016). Table 4 illustrates the numerical steps of the solution.

Normalized correspondences left image (T1*P1)			Normalized correspondences right image (T2*P2)			A matrix in equation 17									SVD f elements
0.74420	0.19085	1	0.00737	0.88884	1	0.00549	0.00141	0.00737	0.66148	0.16964	0.88884	0.74420	0.19085	1	0.02553
1.95370	-0.24045	1	1.23990	0.62191	1	2.42250	-0.29814	1.23990	1.21510	-0.14954	0.62191	1.95370	-0.24045	1	0.04350
0.74433	-0.19961	1	0.06280	0.51715	1	0.04674	-0.01254	0.06280	0.38493	-0.10323	0.51715	0.74433	-0.19961	1	0.50560
2.43150	0.45812	1	1.71390	1.43560	1	4.16740	0.78518	1.71390	3.49070	0.65769	1.43560	2.43150	0.45812	1	-0.07629
2.36390	0.53598	1	1.64520	1.50540	1	3.88920	0.88181	1.64520	3.55870	0.80689	1.50540	2.36390	0.53598	1	0.02066
0.57462	0.06459	1	-0.14162	0.75919	1	-0.08138	-0.00915	-0.14162	0.43625	0.04904	0.75919	0.57462	0.06459	1	-0.28949
1.17040	0.24878	1	0.42785	0.98763	1	0.50074	0.10644	0.42785	1.15590	0.24570	0.98763	1.17040	0.24878	1	-0.44038
3.55820	0.02191	1	2.63810	1.01950	1	9.38680	0.05779	2.63810	3.62750	0.02234	1.01950	3.55820	0.02191	1	0.38847
-0.26779	0.61614	1	-1.11570	1.26750	1	0.29877	-0.68741	-1.11570	-0.33942	0.78095	1.26750	-0.26779	0.61614	1	0.55401
1.60420	-0.49671	1	0.92995	0.33806	1	1.49180	-0.46192	0.92995	0.54230	-0.16792	0.33806	1.60420	-0.49671	1	
2.49980	-0.17701	1	1.71380	0.71749	1	4.28420	-0.30337	1.71380	1.79360	-0.12701	0.71749	2.49980	-0.17701	1	
1.90640	-1.29610	1	1.28700	-0.32102	1	2.45360	-1.66810	1.28700	-0.61199	0.41607	-0.32102	1.90640	-1.29610	1	
2.18160	-0.40172	1	1.45090	0.48718	1	3.16530	-0.58285	1.45090	1.06280	-0.19571	0.48718	2.18160	-0.40172	1	
1.20920	-0.33459	1	0.53847	0.44370	1	0.65114	-0.18017	0.53847	0.53654	-0.14846	0.44370	1.20920	-0.33459	1	
2.39260	-0.66840	1	1.66030	0.27456	1	3.97240	-1.10970	1.66030	0.65692	-0.18352	0.27456	2.39260	-0.66840	1	

F		
0.025525	0.043497	0.505600
-0.076292	0.020655	-0.289490
-0.440380	0.388470	0.554010

Conditioning the Fundamental Matrix to enforce the rank constraint F_conditioned =		
0.02538	0.04329	0.50563
-0.07650	0.02037	-0.28945
-0.44035	0.38851	0.55401

Denormalizing the Fundamental Matrix F = T2'*F_conditioned*T1		
3.566E-08	6.083E-08	4.561E-04
-1.075E-07	2.862E-08	-2.518E-04
-4.489E-04	3.866E-04	3.633E-01

F=F/F(3,3)		
9.81E-08	1.67E-07	1.26E-03
-2.96E-07	7.88E-08	-6.93E-04
-1.24E-03	1.06E-03	1.00E+00

Table 4. The normalized image correspondences and the fundamental matrix derivation.

The final fundamental matrix computed is

$$F = \begin{bmatrix} 9.8137 \times 10^{-8} & 1.6743 \times 10^{-7} & 0.0012553 \\ -2.9586 \times 10^{-7} & 7.8773 \times 10^{-8} & -0.0006930 \\ -0.0012355 & 0.001064 & 1 \end{bmatrix}$$

Furthermore, the epipolar lines of the same corresponding points are calculated as shown in Figure 11 where they indicate a similar solution to one shown using the coplanarity equations.

Figure 11. The image point correspondences and the associated epipolar lines on the stereo images using the fundamental matrix – test 1.

The E matrix is then decomposed into rotation and translation as shown in Table 5.

	ω_{deg}	φ_{deg}	k_{deg}	bY	bZ
Using E	8.7081	-9.5799	6.5027	-1.1751	0.5239

Table 5. Relative orientation parameters using E matrix.

As mentioned in section 2 of the methodology, the distance between 12 image points that are not counted for the computations and the corresponding derived epipolar lines are computed to assess the performance of the three presented techniques (Figure 12).

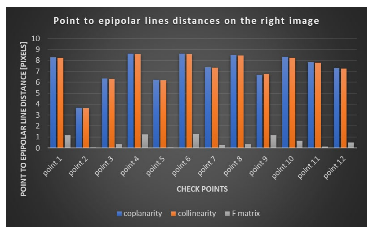

Figure 12. Bar plots of the distance differences– test 1.

To visualize the distance differences, Figure 13 is prepared to show three points to their associated epipolar lines. On the left, the differences are visible of ≅7 pixels using collinearity and coplanarity while around half-pixel on the right using the fundamental matrix.

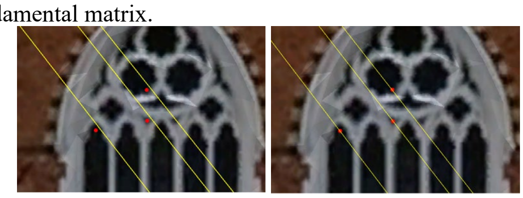

Figure 13. The epipolar lines to points distances are visible using the coplanarity and collinearity methods (left) compared to the distances found using the F matrix (right).

https://doi.org/10.5194/isprs-annals-V-5-2022-25-2022 | © Author(s) 2022. CC BY 4.0 License.

3.2 Second test

The second test is applied on two aerial images taken by the Leica City Mapper camera. The camera parameters are:

- focal length =83 mm
- pixel size= 5.2 µm
- format size= 10336×7788 pixels

The ten image correspondences in pixels are shown in Table 6

	x left	y left	x right	y right
point 1	3556.108	353.049	3531.743	1886.107
point 2	4925.915	2226.046	4899.958	3765.610
point 3	4651.606	1565.489	4626.512	3097.389
point 4	1003.965	1832.618	982.438	3363.495
point 5	1138.639	383.015	1116.408	1916.129
point 6	2285.525	1983.704	2262.830	3514.305
point 7	1738.772	1761.984	1716.222	3293.560
point 8	1787.852	1434.615	1765.327	2966.638
point 9	5127.784	1651.893	5101.955	3189.943
point 10	844.961	994.823	823.095	2526.558

Table 6. Image point correspondences – test 2

A- Using coplanarity equations

Similar to the first test, start the solution using the coplanarity equations given 10 homologous image points. Accordingly, the solution is applied and iterated until converged to insignificant correction values.

The nonlinear least-squares adjustment continues and stops when the corrections are negligible (<10^{-5}) as shown in Figure 14.

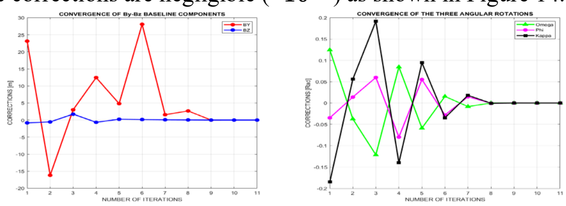

Figure 14. The convergence of the relative orientation parameters corrections using coplanarity in test 2.

Table 7 summarizes the final relative orientation parameters using coplanarity equations.

	ω_{deg}	φ_{deg}	k_{deg}	bY	bZ
Adjusted	-8.816×10^{-2}	1.697×10^{-2}	-1.032×10^{-2}	59.5893	0.2408
σ	$\pm 6.57 \times 10^{-4}$	$\pm 3.85 \times 10^{-4}$	$\pm 9.68 \times 10^{-5}$	± 7.1702	± 0.0440
Initial	0	0	0	1	0

Table 7. Relative orientation parameters using coplanarity – test 2

Furthermore, the epipolar lines of the same corresponding points of Table 1 are calculated using equations 6-10 as shown in Figure 15.

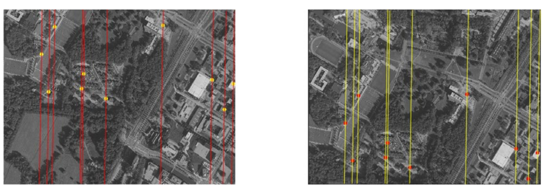

Figure 15. The image point correspondences and the epipolar lines on the stereo aerial images using photogrammetric coplanarity or collinearity equations – test 2.

B- Using collinearity equations

Using collinearity equations, the solution is applied and iterated until converged to insignificant correction values.

The nonlinear least-squares adjustment continues and stops when all the corrections are negligible (<10^{-5}) as shown in Figure 16.

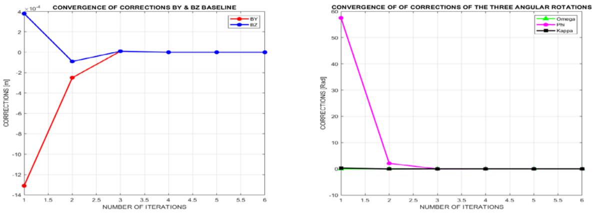

Figure 16. The convergence of the relative orientation parameters corrections using collinearity– test 2.

As logically expected, the same parameter values are estimated using the coplanarity equations as shown in Table 8. Additionally, the epipolar lines of the same corresponding points used to run coplanarity are calculated using equations 11-15 as shown in Figure 15.

	ω_{deg}	φ_{deg}	k_{deg}	bY	bZ
Adjusted	-8.879×10^{-2}	1.719×10^{-2}	-1.032×10^{-2}	59.5204	0.24013
σ	$\pm 6.58 \times 10^{-4}$	$\pm 3.85 \times 10^{-4}$	$\pm 9.70 \times 10^{-5}$	± 7.1675	± 0.0436
Initial	0	0	0	1	0

Table 8. Relative orientation parameters using collinearity– test 2.

C- Using the fundamental matrix

Find the transformation that normalizes the image points. Accordingly, two matrices T1 and T2 will be found:

$$T1 = \begin{bmatrix} 5.9961 \times 10^{-4} & & -0.78153 \\ & 5.9961 \times 10^{-4} & -1.4297 \\ & & 1 \end{bmatrix}, \quad T2 = \begin{bmatrix} 6.1169 \times 10^{-4} & & -1.1049 \\ & 6.1169 \times 10^{-4} & -1.7671 \\ & & 1 \end{bmatrix}$$

Then, using Linear Least Squares, solving the fundamental matrix is

$$F = \begin{bmatrix} 4.4806 \times 10^{-11} & -9.0602 \times 10^{-8} & -0.0029758 \\ 9.0433 \times 10^{-8} & -2.2347 \times 10^{-12} & -0.00069393 \\ 0.0028345 & 0.00069267 & 1 \end{bmatrix}$$

Furthermore, the epipolar lines of the same corresponding points are calculated as shown in Figure 17 where they indicate a similar solution to one shown using the coplanarity equations.

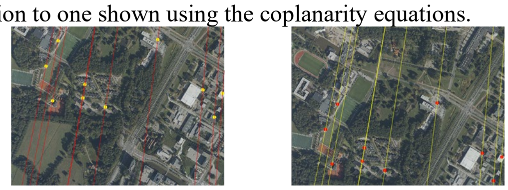

Figure 17. The image point correspondences and the associated epipolar lines on the stereo images using the fundamental matrix in test 2.

As noted in section 2 of the methodology, the distance between 39 image points that are not counted for the computations and the corresponding derived epipolar lines are computed to assess the performance of the three presented techniques (Figure 18).

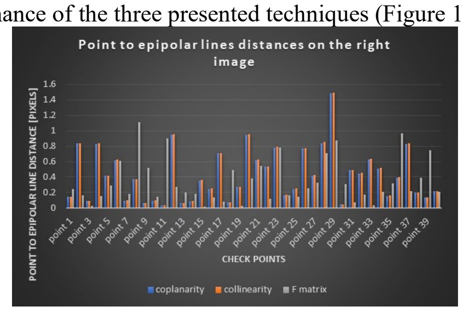

Figure 18. Bar plots of the distance differences – test 2

To visualize the distance differences, Figure 19 is prepared to show three points to their associated epipolar lines. On the left, the differences are visible of ≅ 0.46 pixels using collinearity and coplanarity while ≅ 0.37 on the right using the fundamental matrix.

a) b)

Figure 19. The epipolar lines to points distances – test 2. a) using the coplanarity and collinearity methods. b) using the F matrix (right).

4. DISCUSSION AND CONCLUSION

This paper introduced a clear computational description for the new learners of the photogrammetry field in terms of computing the epipolar geometry for a stereo pair of images. Three mathematical techniques are explained using coplanarity, collinearity, and fundamental matrix.

From the experiments applied in section 3, we noticed that collinearity and coplanarity are identical to give similar relative orientation results. Both methods are solved using nonlinear least-squares adjustment and must start from reliable initial values. This is can be considered as a disadvantage compared to the direct solution following a computer vision-based solution using the fundamental matrix. Furthermore, using the fundamental matrix approach, we didn’t have to know the camera parameters to solve the epipolar geometry.

In terms of minimum data required to solve the relative orientation problem, at least 5 matching image points are required for running the coplanarity and collinearity equations while at least 7 points are required to estimate the fundamental matrix. However, using the current keypoints detectors can result in thousands of valid points which means there will be always a wealth of points to run the orientation problem. This is can be useful to select the points spreading all over the images to ensure a geometrical robust solution when possible.

On other hand, from both coplanarity and collinearity models, it is feasible to estimate the uncertainty in the computed parameters as shown in Tables 2 and 3 in the first test and apply further statistical analysis for blunder detection and analyze the goodness of the least-squares adjustment of relative orientation. What is noticed about collinearity is the sensitivity to the proper initial values compared to coplanarity and this is can be related to the high number of parameters to be estimated when running collinearity which also requires the estimation of the object point XYZ coordinates.

As shown in Figure 12 of the first test, the Fundamental matrix resulted in a higher accuracy to represent the epipolar geometry compared to the coplanarity and collinearity equations. This result indicated that the presented technique to reconstruct the epipolar geometry using coplanarity and collinearity models is more vulnerable to rotational variations. As a final conclusion, if the uncertainty indexes are required in the image orientation, then collinearity and coplanarity are the preferred methods. On the other hand, when the aim is to have a robust direct solution regardless of the prior knowledge about the camera parameters then the fundamental matrix is the preferred solution. Future work can be applied to illustrate more comparisons about multiview stereo vision and the bundle adjustment in both fields of photogrammetry and computer vision.

REFERENCES

Alsadik, B., 2019. Adjustment Models in 3D Geomatics and Computational Geophysics: With MATLAB Examples. Elsevier Science.

Barrile, V., Bilotta, G., Pozzoli, A., 2017. Comparison between innovative techniques of photogrammetry. ITM Web of Conferences 9, 03010.

El-Ashmawy, K.L.A., 2015. A comparison study between collinearity condition, coplanarity condition, and direct linear transformation (DLT) method for camera exterior orientation parameters determination. Geodesy and Cartography 41, 66-73.

Forstner, W., 2002. Computer vision and photogrammetry, Mutual questions, Geometry, statistics and cognition, International Symposium Photogrammetry meets Geoinformatics.

Hartley, R., Zisserman, A., 2003. Multiple View Geometry in Computer Vision. Cambridge University Press.

Hongdong, L., Hartley, R., 2006. Five-Point Motion Estimation Made Easy, 18th International Conference on Pattern Recognition (ICPR'06), pp. 630-633.

Kim, J.-I., Kim, T., 2016. Comparison of Computer Vision and Photogrammetric Approaches for Epipolar Resampling of Image Sequence. Sensors 16, 412.

Liu, G., Klette, R., Rosenhahn, B., 2006. Collinearity and Coplanarity Constraints for Structure from Motion. , Advances in Image and Video Technology. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg.

Luhmann, T., Robson, S., Kyle, S. and Boehm, J., 2014. Close-range photogrammetry and 3d imaging, 2nd ed. De Gruyter (De Gruyter textbook), Berlin.

Nister, D., 2004. An efficient solution to the five-point relative pose problem. IEEE Transactions on Pattern Analysis and Machine Intelligence 26, 756-770.

Salma, C.C., 1980. Manual of Photogrammetry, 4th ed. American Society of Photogrammetry, Falls Church, Virginia.

Sarkar, R., 2016. ECE 661 Homework 10, Computer Vision. Purdue University.

Schindler, K., 2015. Mathematical Foundations of Photogrammetry, in: Freeden, W., Nashed, M.Z., Sonar, T. (Eds.), Handbook of Geomathematics. Springer Berlin Heidelberg, Berlin, Heidelberg, pp. 3087-3103.

Szeliski, R., 2011. Structure from motion, in: Szeliski, R. (Ed.), Computer Vision: Algorithms and Applications. Springer London, London, pp. 303-334.

Weisstein, E.W., 1996. Wolfram MathWorld - Point-Line Distance--2-Dimensional. the Web's Most Extensive Mathematics Resource. Champaign, IL :Wolfram Research.

Wolf, P., DeWitt, B., 2000. Elements of Photogrammetry with Applications in GIS third edition ed. McGraw-Hill