Available online at www.ejournals.uofk.edu
UofKEJ Vol. 10 Issue 2 pp. 50-55 (August 2020)

UNIVERSITY OF KHARTOUM ENGINEERING JOURNAL (UofKEJ)

On The Effects of Equal Weights on the Derived Parameters From Least Squares Solution

Gamal H. Seedahmed
Dept. of Surveying Engineering, Faculty of Engineering, University of Khartoum
(E-mail: Gamal.Seedahmed@gmail.com)

Abstract: Equal weighting is a general strategy in the least squares solutions to reflect the equal contribution of observations that were obtained, for example, by identical measurement systems or similar measurement procedures or algorithms. This type of weighting can be imposed either implicitly or explicitly. Implicit weighting takes the form of an identity weight matrix while explicit weighting is imposed by a weight matrix of equal and known variance value of the observations. Through theoretical and numerical demonstrations, this paper shows that equal weights do not affect the estimated parameters and the residuals in the least squares solution. Moreover, for a relatively large set of observations, the estimated variance component converges to the variance of the original observations in the case of the implicit weighting; and it converges to a value that is very close to one in the case of explicit weighting. In addition, the posterior variance-covariance or dispersion matrices in the implicit and explicit cases are very close to each other after the adjustment. In this study, Monte Carlo simulation was used to generate numerical values of random noise from a normal distribution. This random noise was added to the coordinates of a straight-line for practical evaluation of the proposed arguments.

Keywords: *Equal Weights, Least Squares Solutions, Implicit Weighting, Explicit Weighting, Monte Carlo Simulation.*

1. INTRODUCTION

Least squares solution is a fundamental approach and tool for parameters estimation in geomatics [1] and other fields [2]. Weighted observations in terms of their variances are typically used to reflect the quality of different measurement technologies as well as different measurement procedures. On the other hand, equal weighting is a general strategy in the least squares solutions to reflect the equal contribution of observations that were obtained, for example, by an identical measurement's system or similar measurement procedures such as image matching and feature extraction in digital photogrammetry [3]. It is important to state that equal weighting can be obtained by transformation in terms of Cholesky factorization [4] in which the weight matrix and observations will be transformed into uncorrelated measurements and have equal variances. These form of observations are called homoscedastic observations.

In this paper we are addressing the observations in their original form and without any type of transformation and the concept of equal weighting will be imposed in two different ways. In particular, equal weighting will be imposed either implicitly or explicitly in the target function of the least squares minimization. Implicit weighting takes the form of an identity weight matrix while explicit weighting is imposed by a weight matrix of equal and known variance value of the observations.

Although equal weighting strategy of the observations is a well-known practice in geomatics and surveying [5], it was not treated with the depth that will be provided in this paper. Monte Carlo simulation will be used in this research to add the random noise to the observations. In general, Monte Carlo simulation refers to any simulation that encompasses the use of random numbers [6 and 7]. Monte Carlo simulation is an easy and inexpensive approach to generate and develop control experiments in the broad context of statistical modeling. In particular, it will enable us to develop a detailed understanding of the effects of randomness in the forward and backward mode of the solution. In other words, we will be able to conduct a detailed process of reverse engineering on different aspects of the effects of randomness within the framework of statistical modeling. In particular and in the context of this paper, the forward solution refers to the addition of known random noise to the observations; and the backward solution refers to the least squares modeling to recover or retrieve the added noise to the observations in terms of prediction and to estimate the parameters of the functional model that expresses the relationship between the inputs and outputs.

To conduct a Monte Carlo-based experiment, we need a statistical model to represent the assumed population, a set of statistical parameters of the particular experiment, and a way to generate the random numbers using a computer.

This paper is organized as follows. Section two presents the mathematical proof and the methodology. Section three presents the inputs of four test cases that will be used to evaluate the specific aspect of equal weight on the derived parameters from the least squares solution. Section four provides the results and analysis for the test cases outlined in section three. Section five concludes the paper with some recommendations.

2. Mathematical Proofs and Methodology

The mathematical proof of this work will be based on Gauss-Markov Model (GMM), which can be stated as follows:

$$Y = A\xi + e \quad (1.a)$$

If we neglect the e vector in equation (1.a), then the GMM can be approximated as follows:

$$Y \cong A\xi \quad (1.b)$$

The approximation in equation (1.b) reflects the inconsistency between the two sides of the equation; and this is due to the randomness in the observations. A major assumption was made that the functional model on the right hand side of equation (1.b) is an ideal representation of the observations or measurements of the left hand side in the absence of the randomness shown in equation (1.a).

Where:

Y: is the vector of observations.

A: Design matrix.

ξ: Vector of unknown parameters.

e: Vector of true random errors.

GMM expresses a linear relationship between the observations and the unknown parameters of the model under investigation, which generally follows after a linearization of physical, mathematical, or a geometrical relationship.

$$E(A\xi) = E(Y) \quad (2)$$

$$D(Y) = D(e) = \sigma_o^2 P^{-1} \quad (3.a)$$

$$E(e) = 0 \quad (3.b)$$

Where:

E: Expectation operator.

D: Dispersion operator, which can also be called the variance-covariance matrix of the observations.

σ_o^2: Variance of unit weight.

P: Weight matrix of uncorrelated observations, which can also be written as follows for uncorrelated observations:

$$P = \begin{pmatrix} \frac{1}{\sigma^2} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \frac{1}{\sigma^2} \end{pmatrix} = \frac{1}{\sigma^2}\begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix} \quad (4)$$

Where:

σ^2: Given variance of the observations.

Equation (3.a) represents the explicit case of imposing the strategy of equal weighting in which the weight matrix will be constructed from a direct knowledge of the variance that will be associated with the given direct observations such as the coordinates of a straight-line. On the other hand, the implicit weighting will be imposed by the implicit knowledge of equal weighting using the following simple equation of the variance dispersion:

$$D(Y) = D(e) = \sigma_o^2 I_{n\,x\,n} \quad (5)$$

Where:

I: is an n x n identity matrix.

According to equations (3.a and 5), the target functions for the explicit and implicit least squares solutions are as follows:

$$Explicit\ Case:\quad e^T P e + 2\lambda^T(Y - A\xi - e) = min_{\lambda,e,\xi} \quad (6)$$

$$Implicit\ Case:\ e^T I_{n\,x\,n} e + 2\lambda^T(Y - A\xi - e) = min_{\lambda,e,\xi} \quad (7)$$

Where:

λ: Lagrange's multiplier.

The Lagrange's multiplier provides a very elegant mechanism to solve constrained equations [8, 9, 10, 11, and 12] as the ones shown in equations (6 and 7). It should be noted that the only difference between equations (6 and 7) is the appearance and disappearance of the weight matrix P.

The solution vector for the unknown parameters (ξ) of the implicit case shown in equation 7, in which the weight matrix is the identity matrix I, is:

$$\hat{\xi} = (A^T I A)^{-1} A^T I Y = (A^T A)^{-1} A^T Y \quad (8.a)$$

Equation (8.a) can be manipulated as follows:

$$A^T(Y - A\hat{\xi}) = A^T\big(Y - \hat{Y}\big) = A^T\tilde{e} = 0 \quad (8.b)$$

Where:

$\tilde{e}$: is the residuals vector, which will be restated in equation (11)

$\hat{Y}$: are the estimated and consistent observations vector.

Equation (8.b) reveals two key facts:

- The least squares solution transforms the inconsistent set of linear equations shown in (1.b) to consistent equations by replacing Y by $\hat{Y}$.
- The residuals vector is orthogonal to the column space of the design matrix A. In other words, these two entities ($\tilde{e}$ and A) are not correlated and the least squares solution has the ability to predict the implicit or hidden randomness or residuals in the observations. This is equally true in this work for the implicit and explicit weighting strategy since the residuals are independent of the weighting scheme. Moreover, a unique solution for the residuals will be obtained if the design matrix A has a full rank. This fact is very critical since it provides the basis to approximate the unknown variance of equally weighted observations with a quadratic or squared term of the predicted residuals.

The solution vector for the unknown parameters (ξ) of the explicit case shown in equation 6, in which the weight matrix is P:

$$\hat{\xi} = (A^T P A)^{-1} A^T P Y \qquad (9)$$

As shown in equation 4, the information content of the *P* matrix is a diagonal matrix with equal weight, which can be extracted as common value and the *P* matrix will be reduced to the identity matrix as follows:

$$\hat{\xi} = (\sigma^2)(A^T I A)^{-1} \left(\frac{1}{\sigma^2}\right) A^T I Y = (A^T A)^{-1} A^T Y \qquad (10)$$

It is very evident that the variance will be cancelled out in equation (10) and this equation will be reduced to the same form of equation (8.a). Therefore, regardless of the implicit or explicit form of the least squares solution, the weight matrix does not has any impact on the estimated parameters ($\hat{\xi}$).

The residuals will be computed as follows:

$$\tilde{e} = Y - A\hat{\xi} \qquad (11)$$

In light of equations (8.a and 10), the predicted residuals do not depend on the weight matrix.

The estimated variance of unit weight, which is also called the variance component or the reference variance will be computed as follows for the explicit case:

$$\widehat{\sigma}_o^2 = \frac{\tilde{e}^T P \tilde{e}}{r} = \left(\frac{1}{\sigma^2}\right) x \left(\frac{1}{r}\right)\left(\sum_{i=1}^{i=n} \tilde{e}_i^2\right) \qquad (12)$$

Where:

$\widehat{\sigma}_o^2$: Estimated variance of unit weight or variance component after the adjustment process.

r: redundancy number, which is the difference between the number of equations and the number of unknowns.

It should be noted that the weight matrix (*P*) in equation (12) is replaced by the multiplication of the identity matrix and the inverse of the given variance of the observations. By doing this factorization of the weight matrix, we are preparing the ground for further simplification of equation (12).

In light of the analysis shown for equation (8.b), the square values of the residuals shown in equation (12) can be written and approximated as a summation of equal variance as follows:

$$\sum_{i=1}^{i=n} \tilde{e}_i^2 \approx \sum_{i=1}^{i=n}(\tilde{e}_1^2 + \tilde{e}_2^2 + \cdots + \tilde{e}_n^2) \approx \sum_{i=1}^{i=n}(\sigma^2 + \sigma^2 + \cdots + \sigma^2) \cong n\sigma^2 \qquad (13)$$

Now, we need to plug or insert the approximation shown in equation (13) into equation (12):

$$\widehat{\sigma}_o^2 = \frac{\tilde{e}^T P \tilde{e}}{r} = \left(\frac{1}{\sigma^2}\right) x \left(\frac{1}{r}\right)\left(\sum_{i=1}^{i=n} \tilde{e}_i^2\right) \cong \left(\frac{1}{\sigma^2}\right)\left(\frac{n\sigma^2}{r}\right) \qquad (14)$$

For a relatively large number of observations we get:

$$r \cong n \qquad (15)$$

In light of equations (13) and (15), equation (12) can be approximated by the following new equation:

$$\hat{\sigma}_o^2 = \frac{\tilde{e}^T P \tilde{e}}{r} = \left(\frac{1}{\sigma^2}\right) x \left(\frac{1}{r}\right)\left(\left(\sum_{i=1}^{i=n} \tilde{e}_i^2\right) \cong \left(\frac{1}{\sigma^2}\right)\left(\frac{n\sigma^2}{r}\right)\right) \cong 1 \qquad (16)$$

The dispersion matrix for the explicit case after the plugging of the estimated variance component shown in equation (16) can be written as follows:

$$\widehat{\widehat{D\langle\xi\rangle}} = \hat{\sigma}_o^2 (A^T P A)^{-1} = \hat{\sigma}_o^2\, x\, \sigma^2 (A^T I A)^{-1} \cong 1\, x\, \sigma^2 (A^T I A)^{-1} \cong \sigma^2 (A^T I A)^{-1} \qquad (17)$$

By using the approximate results shown in equations (13) and (15), we will get the following results for the variance component and the dispersion matrix for the implicit case:

$$\widehat{\sigma}_o^2 = \frac{\tilde{e}^T I_{n\,x\,n} \tilde{e}}{r} = \left(\frac{1}{r}\right)\left(\sum_{i=1}^{i=n} \tilde{e}_i^2 \cong \left(\frac{n\sigma^2}{r}\right)\right) \cong \sigma^2 \qquad (18)$$

Equation (18) can be used to estimate a good approximation for the unknown variance of the observations. The dispersion matrix for the implicit case can be written as follows:

$$\widehat{\widehat{D\langle\xi\rangle}} = \hat{\sigma}_o^2 (A^T I A)^{-1} \cong \sigma^2 (A^T I A)^{-1} \qquad (19)$$

It is very important to note that equations (17) and (19) converged to a similar approximate solution for the dispersion matrix after the adjustment for the explicit and implicit cases. In other words, this similarity is only available after the multiplication with the estimated variance component values. In the explicit case, the effect of the estimated variance component is very minor or negligible and this is due to the multiplication by its value that is closer to one. On the other hand and for the implicit case shown in equation (19), the estimated variance component approximately imposed the unknown variance of the observations on the dispersion or variance-covariance matrix. Equations (17) and (19) show a very interesting interplay on how the notion of equal weights impacted the dispersion matrix after the estimation of the variance component.

The methodology for evaluating this research work will be based on the following steps:

- As stated before, the equation of the straight-line will be used to test the formulation and the theoretical insights of equal weight. Therefore, the first step will be to generate a set of points that belongs to a straight-line. The following equation will be used for the straight-line representation:

$$y = mx + c \qquad (20)$$

x, y: 2D coordinates of the straight-line.

m: slope.

c: y-intercept

- Generate a set of random noise that belongs to the normal distribution. In particular, a random noise with a zero mean and a specific standard deviation value will be generated.
- Add the random noise to the coordinates of the straight-line. In fact, two situations of random noise will be tested in this paper. The first situation is concerned with the addition of random noise to the *y-coordinates*; and in the second one to both sets of the coordinates.

- Use the presented set of equations to evaluate the effects of equal weight. As stated, in this work we distinguished between explicit and implicit weighting. In the explicit case the weight matrix of equal and known variance was used directly in the formulation of the least squares solution as shown in equation (6). On the other hand, in the implicit case the weight matrix used the generic or the believed assumption of equal variance in the form of identity matrix in the formulation of the least squares solution as shown in equation (7). In other words, there is no knowledge of the actual value of the variance of the observations and the proposed work can be used to estimate a good approximation for its value.

3. Test Cases

As shown in Table 1, four test cases will be used to demonstrate the effects of equal weights on the derived parameters from the least squares solution. Each case will be specified by four parameters, namely, a case number, a standard deviation, number of points, and the specification of the corrupted coordinates. Case 1 and 2 share the same level of random noise but they differ in the number of points. Case number 3 shares the same number of points with case number 2 but it differs in the random noise. Case number 4 shares the same number of points and random noise with case number 2 but it differs in the effect of random noise since it impacts the two coordinates. The last case is very important since it shows a common scenario of least squares solution in which both coordinates are random; and the ordinary solution of least squares does not capture the full effects or interaction between random errors distribution and the derived quantities from least squares solution.

A straight-line with a slope of 1.53986 and an intercept of 50 units will be used in all cases shown in Table 1. These two parameters of the straight-line will be used to generate ideal measurements for the 2D coordinates that belong to this line. In other words, the vectors basis of the design matrix shown in equation (1.a) truly express the functional relationship between the two sides of the equation in the absence of the random noise. In direct statement, we are dealing with controlled experiments. Fig. 1 shows a plot for an example of the coordinates for the straight-line coordinates that will be used for testing the proposed work. These coordinates will be corrupted with random noise that will be generated from normal distribution. Fig. 2 shows a histogram for a random noise that was generated from a normal distribution with a zero mean and ± 0.05 standard deviation.

4. Results and Analysis

Tables 2 and 3 show the results of the first case in which the estimated values of the line parameters are identical (see Table 2) in the explicit and the implicit cases. In other words, the experimental findings confirm the theoretical derivation shown in equations (8.a and 10). Moreover, Table 2 shows the estimated values for the variance components for the explicit and implicit cases in which the value for the explicit case is very close to one (1.03722) and the value for the implicit case is very close to $\pm$0.05 ($\pm$0.0509). Once again, this finding confirms the theoretical proof shown in equations (16) and (18). Table 3 shows the two dispersion matrices for the explicit and implicit cases and both of them are very close or even identical to each other. In other words, the dispersion matrix in the explicit and implicit case converge to similar values as predicted by equations (17) and (19). Similar empirical and theoretical results were obtained for the test cases number 2 and 3 and they were shown in Tables 4, 5, 6, and 7. In test case number 4, the random noise was added to the two coordinates of the straight-line. Recall that this is not the situation for the other three cases shown in Table 1 in which the random noise was added only to the *Y*-coordinates. As stated, case number 4 uses the same variance and the same number of points that were used in test case number 2. Indeed, the estimated line parameters are very identical in the explicit and implicit cases because they were not sensitive to the appearance or disappearance of the weight matrix (see Table 8).

At no surprise, the developed theory of this research does not account for the variances at the explicit and implicit weighting. Moreover, it gives very optimistic values for the dispersion matrices by one-order of magnitude for the variances of the line parameters (see Table 9).

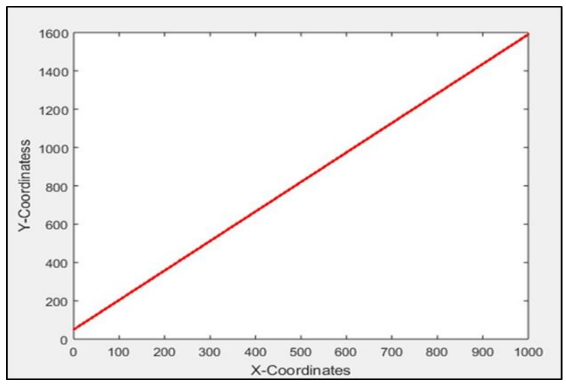

Fig .1. Plot of an example for a straight-line coordinates.

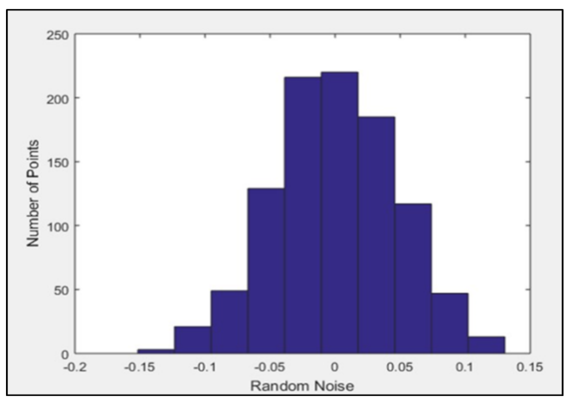

Fig .2. A histogram of random noise from normal distribution.

Table 1. Test cases.

Case Number	Standard Deviation	Number of Points	Corrupted Coordinates
1	±0.05	500	Y
2	±0.05	1000	Y
3	±0.07	1000	Y
4	±0.05	1000	X & Y

Table 2: Results for case number 1.

Parameter	Explicit Case	Implicit Case
$\hat{m}$	1.53989	1.53989
$\hat{c}$	49.99636	49.99636
$\hat{\sigma}_o^2$	1.03722	0.00259, STD=± 0.0509

Table 3. The two dispersion matrices for case number 1.

Weighting Type	**Dispersion Matrix**
Explicit	2.489e-10 -6.235e-08 -6.235e-08 2.080e-05
Implicit	2.489e-10 -6.235e-08 -6.235e-08 2.080e-05

Table 4. Results for case number 2.

Parameter	**Explicit Case**	**Implicit Case**
$\hat{m}$	1.53986	1.53986
$\hat{c}$	49.99850	49.99850
$\hat{\sigma}_o^2$	1.0792	0.00269, STD=±0.0519

Table 5. The two dispersion matrices for case number 2.

Weighting Type	**Dispersion Matrix**
Explicit	3.237e-11 -1.620e-08 -1.620e-08 1.080e-05
Implicit	3.237e-11 -1.620e-08 -1.620e-08 1.080e-05

Table 6. Results for case number 3.

Parameter	**Explicit Case**	**Implicit Case**
$\hat{m}$	1.53986	1.53986
$\hat{c}$	49.99837	49.99837
$\hat{\sigma}_o^2$	0.94093	0.00461 STD=± 0.0679

Table 7. The two dispersion matrices for case number 3.

Weighting Type	**Dispersion Matrix**
Explicit	5.532e-11 -2.769e-08 -2.769e-08 1.846e-05
Implicit	5.532e-11 -2.769e-08 -2.769e-08 1.846e-05

Table 8. Results for case number 4.

Parameter	**Explicit Case**	**Implicit Case**
$\hat{m}$	1.53986	1.53986
$\hat{c}$	50.00087	50.00087
$\hat{\sigma}_o^2$	0.30156	0.00075 STD=±0.0274

Table 9. The two dispersion matrices for case number 4.

Weighting Type	**Dispersion Matrix**
Explicit	9.046e-12 -4.527e-09 -4.527e-09 3.020e-06
Implicit	9.046e-12 -4.527e-09 -4.527e-09 3.020e-06

5. Conclusions and Recommendations

This paper presents a theoretical proof for the effect of equal weighting strategy on the derived parameter from a least squares solution. We distinguished between the explicit and implicit strategies for introducing the weight matrix in the formulation of the least squares solution. In the explicit case, the weight matrix uses the known variance of the observations along its diagonal. On the other hand, the implicit case replaces the unknown variance of the observation by an identity matrix for weighting. As shown in the sequel of this paper, the distinction between implicit and explicit cases of equal weights is irrelevant to the formulation of the target function of least squares optimization. In particular, the implicit weighting will account for both cases. In other words, the target function is unresponsive to the explicit case and intuitively compatible with the notion of equal weights. Interestingly enough and for a relatively large number of observations, the implicit case can be used to estimate a very close approximation for the value of the unknown variance of the observations.

The posterior variance-covariance or the dispersion matrices in the implicit and explicit cases are very close to each other. Experimental results confirmed the theoretical proof for Gauss-Markov Model, which is also known as the ordinary least squares solution. At no surprise, the developed theory cannot account for the correct values for the estimated variance component when the random noise affected the two coordinates in the explicit and implicit cases. Therefore, further work is required to handle the effect of equal weight when the random noise affected both coordinates. In general, the presented approach can be used to evaluate the accuracy of the measuring instruments in geometrics and other fields.

Monte Carlo simulation provides an elegant and exciting mechanism to test the validity of the proposed work in terms of supplying unlimited number of test cases with a chosen level of uncertainty or random noise. Accordingly, a whole spectrum of testing campaign can be carried out for deep understanding and analysis for the different functional models within the framework of adjustment computations on how they interact with the randomness in geomatics and other fields.

REFERENCES

[1] Ogundare, J. (2019). Understanding Least Squares Estimation and Geomatics Data Analysis. John Wiley & Sons, Inc. 696 p.

[2] Kailath, T., Sayed, A., and Hassibi, B. (2000). Linear Estimation. Prentice Hall. 854 p.

[3] Schenk, T. (1999). Digital Photogrammetry. Terra Science. 428 p.

[4] Kock, K. (1999). Parameter Estimation and Hypothesis Testing in Linear Models. Springer. 333 p.

[5] Ghilani, C. and Wolf, P. (2006). Adjustment Computations: Spatial Data Analysis. 4th edition. John Wiley & Sons, Inc. 611 p.

[6] Robert, C. and Casella, G. (2004). Monte Carlo Statistical Methods. 2nd edition. Springer. 645 p.

[7] Martinez, W. and Martinez, A. (2005). Computational Statistics Handbook with MATLAB. 2nd edition. Chapman and Hall/CRC. 767 p.

[8] Seedahmed, G. (2017). On The Integration and Evaluation of Vertical Control Information and Uncertainties in Leveling Networks Using Least Squares Modeling. University of Khartoum Engineering Journal, Vol. 7(2): pp: 25-33.

[9] Seedahmed, G. (2013). Relative Orientation of Stereo-Pair Using 2D Projective Singular Correlation: Applications on Video Images and Scanned-Aerial Photographs. University of Khartoum Engineering Journal, Vol. 3(2): pp: 12-23.

[10] Seedahmed G. (2004). *On The Suitability of Conic Section in A Single-Photo Resection, Camera Calibration, and Photogrammetric Triangulation*. Ph.D. Dissertation, Dept. of Civil & Environmental Engineering & Geodetic Science, The Ohio State University, Columbus, Ohio, USA, 138 p.

[11] Seedahmed, G. and Schenk, T. (2001). Direct Linear Transformation in the Context of Different Scaling Criteria. In Proc. *ASPRS. Ann. Convention*, St. Louis, USA.

[12] Seedahmed, G. and Schenk, T. (2001). Comparative Study of two Approaches of Deriving the Camera Parameters from Direct Linear Transformation. In Proc. *ASPRS. Ann. Convention*, St. Louis, USA.