

NETWORK ORIENTATION MODELS FOR IMAGE-BASED 3D MEASUREMENT

Clive S. Fraser

Department of Geomatics, University of Melbourne, Victoria 3010, Australia
c.fraser@unimelb.edu.au

KEY WORDS: close-range photogrammetry, image orientation, automation, relative orientation, collinearity, projective geometry

ABSTRACT: Network orientation is a fundamental step in image-based 3D measurement. This paper aims to offer insights into network orientation by considering developments in analytical close-range photogrammetry during its evolution from a film-based to digital image-based measurement technology suited to full process automation. A brief mathematical background to orientation models derived via perspective projection and projective geometry is first presented, and the issue of generating initial values for image orientation parameters is highlighted. Developments in network orientation processes for close-range photogrammetry are then reviewed and the potential of utilising linear models from projective geometry in the photogrammetric orientation process is discussed. The state of the art in network orientation for close-range photogrammetry is indicated by way of the discussion.

1. INTRODUCTION

Photogrammetry has been applied to the measurement and modelling of complex heritage sites and architectures for a very long time. In the past two decades or so there has been an evolution in close-range photogrammetric recording from film-based stereo restitution with metric cameras to multi-image network orientation of images from consumer-grade digital cameras, with automatic and semi-automatic 3D feature point measurement. Whereas camera and computing technology have witnessed dramatic change during the life of analytical photogrammetry, the underlying mathematical models for network orientation, which are based on perspective projection, remain essentially unchanged. They have served photogrammetry well, and continue to do so.

The well-known models for relative, absolute and exterior orientation are non-linear, yet they are solved via linear least-squares techniques. These require initial approximate or ‘starting’ values for the parameters. This determination of starting values has had a considerable influence upon the development of different computational schemes for close-range photogrammetry. The network orientation process, however, almost always culminates in a ‘rigorous’ solution, most often via a bundle adjustment.

In the more general domain of image-based 3D measurement, which embraces computer vision as well as robot and machine vision, alternative network orientation approaches have been developed. Some of these have been rigorous from a metric standpoint; most have not. Popular in this context are models for 3D motion capture and scene reconstruction based on projective geometry. These are characterised by elegant mathematics, and also by questions as to the computational stability and reliability of the linear solutions adopted in the recovery of image-to-object geometry. One motivation for adopting the linear solution algorithms of projective geometry is that they circumvent the need to provide the initial values that are required with photogrammetric orientation models based on perspective projection.

This paper aims to offer further insights into the network orientation problem for image-based 3D measurement. A brief

mathematical background to orientation models based on perspective projection and projective geometry is first offered. This is followed by a review of the different network orientation scenarios that have been employed in close-range photogrammetry over the past 25 years, i.e. in the period of development of stand-alone measurement systems designed for a broad range of applications, including heritage recording. The role of projective geometry models is considered in the context of seeking better ways to determine initial values for image orientation parameters. In discussing network orientation approaches, the paper also reports on the state of the art in close-range photogrammetry and vision metrology, which is represented by systems ranging from fully automatic to those with on-line network orientation to support manual measurement of images recorded with off-the-shelf digital cameras.

2. MATHEMATICAL BACKGROUND

2.1 Collinearity Model

The fundamental goal of photogrammetry is to obtain 3D measurements of objects from multiple, overlapping 2D images. In basic terms, what is required is a transformation of coordinates between two Cartesian spaces, image space (x, y, z) and object space (X, Y, Z) . This can be written as a 7-parameter similarity transformation:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \lambda R \begin{pmatrix} X - X^c \\ Y - Y^c \\ Z - Z^c \end{pmatrix} \quad (1)$$

where the rotation matrix R is formed as the product of three 2D rotations through angles which are identified here as azimuth α , elevation ε and roll κ . The translation terms X^c , Y^c and Z^c express the position of the image space coordinate system in the object space system. If the scale factor λ is 1, Eqn. 1 represents a rigid body transformation. If we take into account that the image points are all within a focal plane, then the 3D-to-2D perspective projection follows from Eqn. 1 as

$$\begin{pmatrix} x - x_p + dx \\ y - y_p + dy \\ -c \end{pmatrix} = \lambda R \begin{pmatrix} X - X^c \\ Y - Y^c \\ Z - Z^c \end{pmatrix} = \lambda \begin{pmatrix} X' \\ Y' \\ Z' \end{pmatrix} \quad (2)$$

where c is the principal distance and x_p, y_p the coordinates of the principal point. These three parameters constitute the interior orientation elements of the camera. The terms dx and dy represent the corrections for so-called departures from collinearity, which arise from image distortions. The principal perturbation to image point position is radial lens distortion.

Eqn. 2 can also be cast in a form more familiar to the computer vision community, who utilise methods of projective geometry in their building of 3D object models from imagery:

$$\frac{1}{\lambda} \begin{pmatrix} 1 & 0 & -x_p \\ 0 & 1 & -y_p \\ 0 & 0 & -c \end{pmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = R \begin{pmatrix} 1 & 0 & 0 & -X^c \\ 0 & 1 & 0 & -Y^c \\ 0 & 0 & 1 & -Z^c \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \quad (3)$$

or

$$\frac{1}{\lambda} C \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} = P \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \quad (4)$$

Here, the 3x4 matrix P is termed the projection matrix and C the 'calibration' matrix, even though it typically does not take into account the very significant lens calibration corrections represented by dx and dy in Eqn. 2.

As written, there is no distinction between the homogenous representation of Eqn. 3 and the form more familiar to photogrammetrists (Eqn. 2), except of course for the lens distortion corrections dx and dy . It will be seen, however, that in solving for the image-to-object space transformation there is a marked difference in approach. A simple division of the first and second rows of Eqn. 2 by the third gives rise to the well-known collinearity equations

$$\begin{pmatrix} x - x_p + dx + c X'/Z' \\ y - y_p + dy + c Y'/Z' \end{pmatrix} = 0 \quad (5)$$

The solution of the non-linear equation system generated by bundles of intersecting rays from multiple images is via a linear least-squares approach, which requires linearization of Eqn. 5 to the general form of the photogrammetric bundle adjustment:

$$-v + A_1 \delta_1 + A_2 \delta_2 + w = 0 \quad (6)$$

Here, v is the vector of observational residuals (residuals in image coordinate measurements); A_1 and A_2 are matrices of partial derivatives; w is a discrepancy vector and δ_1 and δ_2 comprise the corrections to starting values for the six exterior orientation parameters ($\alpha, \varepsilon, \kappa, X^c, Y^c, Z^c$) and three object point coordinates (X, Y, Z). It is not the intention here to develop the well-known bundle adjustment formulation any further, as the reader can find this in modern textbooks on photogrammetry (e.g. Krauss, 2000; Mikhail et al., 2001). What is important to this discussion is firstly that the solution to Eqn. 6 is rigorous in a functional and stochastic sense (it is a maximum likelihood solution), and secondly that in order to recover δ_1 and δ_2 appropriate starting values for the parameters are required. These initial values will here be termed \underline{Q} and \underline{X} , where

$$\underline{Q} = (\alpha_1^0, \varepsilon_1^0, \kappa_1^0, X_1^{c0}, Y_1^{c0}, Z_1^{c0}, \dots, Y_m^{c0}, Z_m^{c0})^T$$

and

$$\underline{X} = (X_1^0, Y_1^0, Z_1^0, \dots, X_n^0, Y_n^0, Z_n^0)^T$$

for the m images and n object points involved.

Since the emergence of analytical photogrammetric methods in the 1960s and 1970s, much of the attention given to finding practical solutions for Eqn. 6 has centred upon developing efficient means to determine starting values. As will be seen later, the most common approach has involved separately solving for δ_1 and δ_2 . Under the assumption that a given number of 'control points' (known in XYZ) are available, δ_2 can be suppressed and δ_1 is solved image by image in a series of spatial resections. Having determined the image orientation, δ_1 is suppressed and δ_2 is solved by spatial intersection. With the initial values \underline{Q} and \underline{X} in place, a full bundle adjustment solution follows. More will be said later about the scenarios for exterior orientation.

To non-photogrammetrists, the need to determine initial values seems to have been an impediment to the adoption of rigorous photogrammetric approaches, as was both the explicit requirement for image coordinates referenced to the principal point (base of the perpendicular from the projection centre to the image plane) and the implicit necessity of accounting for calibration corrections such as lens distortion. Thus, projective geometry formulations as represented by Eqn. 3 appeared as potential alternatives because they could be solved in a linear manner. This assumed that the elements p_{ij} of the projection matrix P are linearly independent, which of course they are not if the equivalence of Eqns. 2 and 3 is to hold true. The first of these linear models, which predated early developments in computer vision approaches by a decade, was the well-known direct linear transformation (DLT) which was introduced to close-range photogrammetry by Abdel-Aziz & Karara (1971). This modelled the projective geometry relationship between image coordinates (x', y') of arbitrary scale, orientation and origin, and object space coordinates.

The DLT is generally solved in a two-step process, equivalent to spatial resection and intersection, though a 'bundle adjustment' formulation is possible. Of the 11 parameters involved, only 9 are independent and thus difficulties can be expected with certain configurations of camera stations and object points, the most obvious being that the DLT will not handle situations where the object point array is planar or near planar. As a consequence of the two non-linear constraints implicit in its formulation being ignored, the DLT has a tendency to be numerically unstable, especially in situations of low observational redundancy, say with 10 or less 'known' object points; a minimum of 6 is required. The DLT can accommodate calibration corrections dx and dy , but only where 'true' image coordinates (x, y) are employed, rather than (x', y').

2.2 Coplanarity Model

In any discussion of approaches to orientation of close-range photogrammetric networks it is necessary to consider a second fundamental model which, at least in analog form, has been the basis of stereo photogrammetry for more than half a century. This is the coplanarity model that states that the two intersecting rays to an object point from two images must lie in a single plane, which also includes the baseline vector between the two perspective centres. This plane is called the epipolar

plane. The coplanarity condition can be formulated, again making use of the perspective projection, Eqn. 2, for the case of one image being relatively oriented to a second, as

$$u^T \begin{pmatrix} 0 & b_z & -b_y \\ -b_z & 0 & b_x \\ b_y & -b_x & 0 \end{pmatrix} R_2^T u = 0 \quad (7)$$

where

$$u = (x - x_p + dx \quad y - y_p + dy \quad -c)^T$$

Here, the matrix R_2 describes the rotation of the second image with respect to the first and b_y and b_z are translations. The translation b_x , which lies in the base line, can be assigned an arbitrary value, since scale cannot be recovered from the coplanarity model. Note also, that object space coordinates do not explicitly appear.

Equation 7 can also be recast into homogeneous form, along the same lines as Eqn. 3. For the case of two images from the same camera, the expression becomes

$$\begin{pmatrix} x_1 & y_1 & 1 \end{pmatrix} C^T K R_2^T C \begin{pmatrix} x_2 \\ y_2 \\ 1 \end{pmatrix} = 0 \quad (8)$$

Here, K represents the skew-symmetric matrix in Eqn. 7 and C is again the calibration matrix (without consideration of dx or dy). Further substitution for matrix products leads to

$$\begin{pmatrix} x_1 & y_1 & 1 \end{pmatrix} C^T E C \begin{pmatrix} x_2 \\ y_2 \\ 1 \end{pmatrix} = 0 \quad (9)$$

and

$$\begin{pmatrix} x_1 & y_1 & 1 \end{pmatrix} F \begin{pmatrix} x_2 \\ y_2 \\ 1 \end{pmatrix} = 0 \quad (10)$$

where E is called the essential matrix, and F the fundamental matrix. The distinction between them is the assumption that the interior orientation is known for the essential matrix expression (Hartley & Zisserman, 2000; Faugeras & Luong, 2001). The non-iterative algorithm for relative orientation (RO) via the essential matrix, which has been widely adopted in computer vision, has been attributed to Longuet-Higgins (1981). However, it had already been known within the photogrammetric community for at least two decades, as illustrated by Thompson (1959) and Stefanovic (1972).

Once again, the projective geometry models for RO, which algebraically represent epipolar geometry and centre upon the essential and fundamental matrices, are equivalent to the coplanarity condition, at least when the lens distortion corrections dx and dy are ignored. However, the solution of Eqn. 9 by linear methods, which assume that the elements of E and F are independent, is not the same as solving Eqn. 7 via a Gauss-Markov model. The linear solution for the three rotations and two translations of RO via the essential matrix, which is more appropriate in a photogrammetric context than the fundamental matrix, is generally a two-step process. The elements e_{ij} of E are first solved via the expression

$$A \underline{e} = 0 \quad (11)$$

where

$$A_i = (x_1 x_2 \quad y_1 x_2 \quad x_2 \quad x_1 y_2 \quad y_1 y_2 \quad y_2 \quad x_1 \quad y_1 \quad 1)_i$$

and

$$\underline{e} = (e_{11}, e_{12}, e_{13}, e_{21}, e_{22}, e_{23}, e_{31}, e_{32}, e_{33})^T$$

after which E is decomposed into its constituent component matrices K and R_2 .

In much the same way as was previously described for the collinearity model, the linear least-squares model for the coplanarity condition is given as (Mikhail et al., 2001):

$$Bv + A\delta + w = 0 \quad (12)$$

where A and B are matrices of partial derivatives with respect to the parameters and to image coordinates, respectively; v and w are as previously defined for Eqn. 6; and δ is the vector of corrections to the initial values for the three rotation angles (taken this time as Euler angles ω , φ , κ) and two translations. These initial values are represented by the vector $\underline{\Delta}$:

$$\underline{\Delta} = (\omega^0, \varphi^0, \kappa^0, by^0, bz^0)^T$$

For convergent imaging configurations with arbitrary image orientation (e.g. the camera is ‘rolled’ between portrait and landscape orientation), the determination of appropriate initial values $\underline{\Delta}$ can be very challenging, hence the appeal of linear solutions such as the essential and fundamental matrix approaches. One can read not only of the recovery of RO parameters from a stereo pair of images via the projective geometry approach, but also of the simultaneous determination of certain interior orientation elements, for example the focal lengths associated with each image. Photogrammetrists would state that this is not feasible, at least in practical and metric terms, in spite of the elegance of the mathematics involved in deriving the solutions for E and F and subsequently decomposing these matrices to determine the projection matrix P in Eqn. 3. More will be said on this aspect later. However, it is widely recognised that ‘noisy’ data (read redundant observations and real measurements) can lead to numerically unstable solutions for the essential and fundamental matrices, and consequently to unreliable results.

In the two-step solution process mentioned above, E (or F) is typically determined through either a homogenous linear equation solution with normalised coordinates for the 8 or more points involved (Hartley, 1997), a singular value decomposition or a RANSAC approach (Fischler & Bolles, 1981) in cases where there is an abundance of corresponding point pairs available. The rotation matrix and translations are then recovered by singular value decomposition (Hartley, 1992), which for the essential matrix of Rank 2 should in theory yield two equal singular values and a third which is zero.

The projective geometry approaches are considered in this discussion not because they present a potential alternative to the collinearity and coplanarity models of photogrammetric orientation, but more because they offer at first sight possible practical approaches to the determination of initial values $\underline{\Delta}$, \underline{Q} and \underline{X} . The implementation of such an approach has been reported by Roth (2004). One must keep in mind, however, the difficulties associated with the reliable recovery of RO via projective geometry approaches in cases of ‘difficult’ (read highly convergent) geometry. Moreover, like the DLT, these approaches do not work in instances of a near planar array of object points and they can yield ambiguous solutions. The words of Horn (1990a) are noteworthy in this regard: “Overall, it seems that the two-step approach to relative orientation,

where one first determines an essential matrix, is the source of both limitations and confusion". Alternative RO algorithms employing closed-form initial parameter determination are available (e.g. Horn 1990b).

It is useful, having reviewed the mathematical approaches to the determination of relative and exterior orientation, to now consider the developments in network orientation that have occurred over the past three decades in close-range photogrammetry. The common endpoint of the orientation process for metric applications is - and should be - the bundle adjustment model of Eqn. 6. Critical to implementation of this general model, however, is the generation of initial values. It will be seen how this aspect has greatly influenced practical implementation of analytical photogrammetry, and how the emergence of digital cameras has given impetus to the development of new approaches for multi-image network orientation.

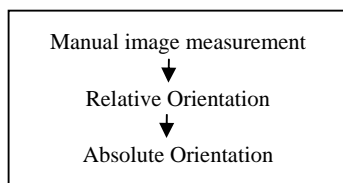
3. NETWORK ORIENTATION SCENARIOS

3.1 Range of Camera Station Configurations

Shown in Figs. 1 to 4 are examples of camera station configurations encountered in close-range photogrammetry. The first exemplifies a 'stereo' geometry, implying low convergence or near parallel optical axes. Figs. 2 and 3 constitute typical convergent configurations, which exemplify strong network geometry. Finally, Fig. 4 represents a photogrammetrically challenging network geometry where the camera stations and most object points are near to being in a single plane. The last case is unfortunately very representative of the network geometry encountered in traffic accident reconstruction, which is a fast growing applications area for digital close-range photogrammetry (Fraser et al., 2005). Reference will be made to these network configurations in the following discussion.

3.2 The Traditional Approach

As analytical photogrammetry evolved, the 'traditional' orientation scenario of analog stereo restitution remained popular. This was a two-step process, appropriate for 2-image configurations:

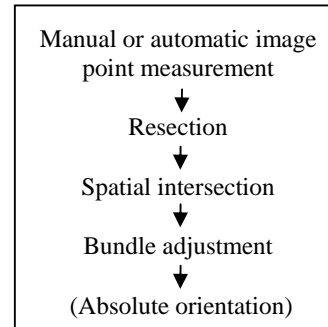


The RO was via the coplanarity model (Eqn. 7) or, less frequently, the collinearity model (Eqn. 5). The absolute orientation (AO) was performed with a 3D similarity transformation (Eqn. 1). Thus, in the present context, the first issue typically concerned how to determine initial values $\underline{\Delta}$. Fortunately, this was relatively straightforward for stereo geometry. Two rotations could be assigned an initial value of zero, and the relative rotation about the optical axis could be estimated from the image point distribution. The initial values for translation were most often taken as zero. For AO, closed-form and quasi least-squares solutions for 3D similarity transformation are well known, and thus computing rigorous

AO via a linear least-squares solution to Eqn. 1 does not pose any practical difficulties. Networks exemplified by Fig. 1 could therefore be oriented reliably, with certain constraints upon the degree of convergence between the two optical axes.

3.3 Early Days with Multi-Image Networks

Given that the approach above was effectively limited to two-image stereo networks, and that stereo geometry is not optimal from an accuracy standpoint, an alternative was sought to accommodate the convergent multi-image geometry shown in Figs. 2 and 3. This gave rise to a second orientation scenario, as follows:



A number of object points (minimum of 4) were assigned preliminary XYZ coordinates (measured or arbitrary), and from these two or more images would be resected via closed-form resection (e.g. Fischler & Bolles, 1981). Thus, \underline{Q} was established for these images. Spatial intersection would follow to provide the XYZ coordinates of object points (\underline{X}). From the new object point coordinates, further images were resected and further points intersected until initial values were established for all parameters. Bundle adjustment then followed to refine the approximate values. This process suited the sequential, monoscopic measurement of images. It had two main drawbacks - which were not viewed as such at the time - namely that initial XYZ coordinate values for at least 4 points were needed and a careful labelling of image points was required to ensure correct correspondences between images.

This approach was adopted in the 70s and 80s for industrial photogrammetry systems, and it remains in common use today. At the same time the computer vision community were engaged in popularising the essential matrix approach, albeit two images at a time. Why, one might ask, were the methods of photogrammetric orientation not adopted in computer vision? Of no doubt many contributing factors, four come to mind:

- i. There was no desire whatever to get involved with manual point labelling; correspondences were to be determined automatically, with a percentage of these accepted as being potentially erroneous.
- ii. The need to assign object point coordinates and determine initial values was to be avoided.
- iii. There was a preference to work with pixel coordinates and to ignore lens calibration.
- iv. Metrically accurate results were not being sought.

Nevertheless, some developments in computer vision were curious given the then state of the art in photogrammetry. For example, the camera calibration approach of Tsai (1987) required not only the provision of an object point array with known XYZ coordinates, but also a multi-stage process

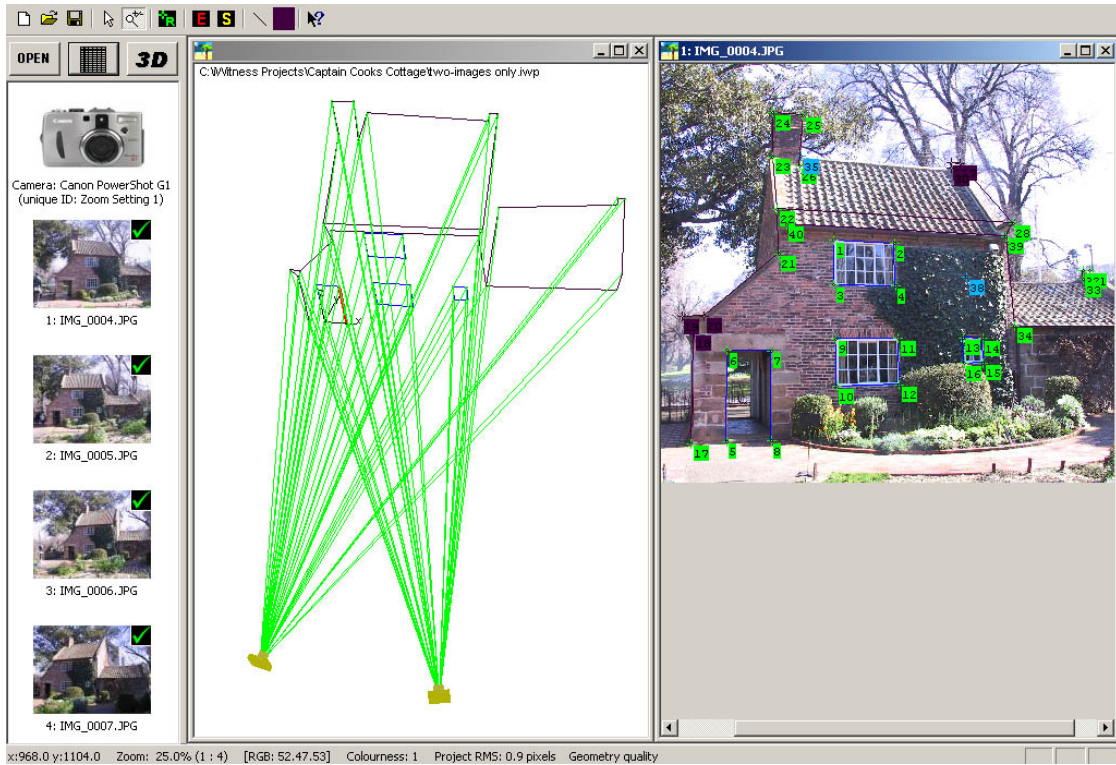


Figure 1: Stereo image geometry with low convergence.

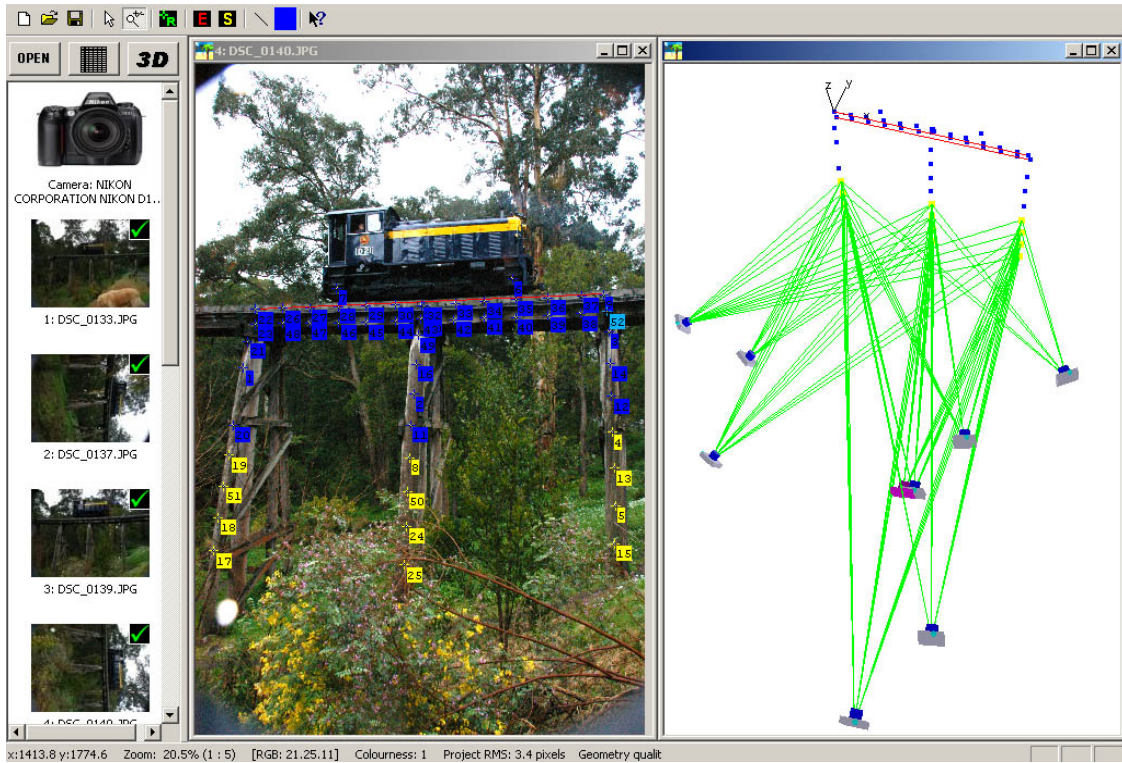


Figure 2: Convergent multi-image network and near-planar array.

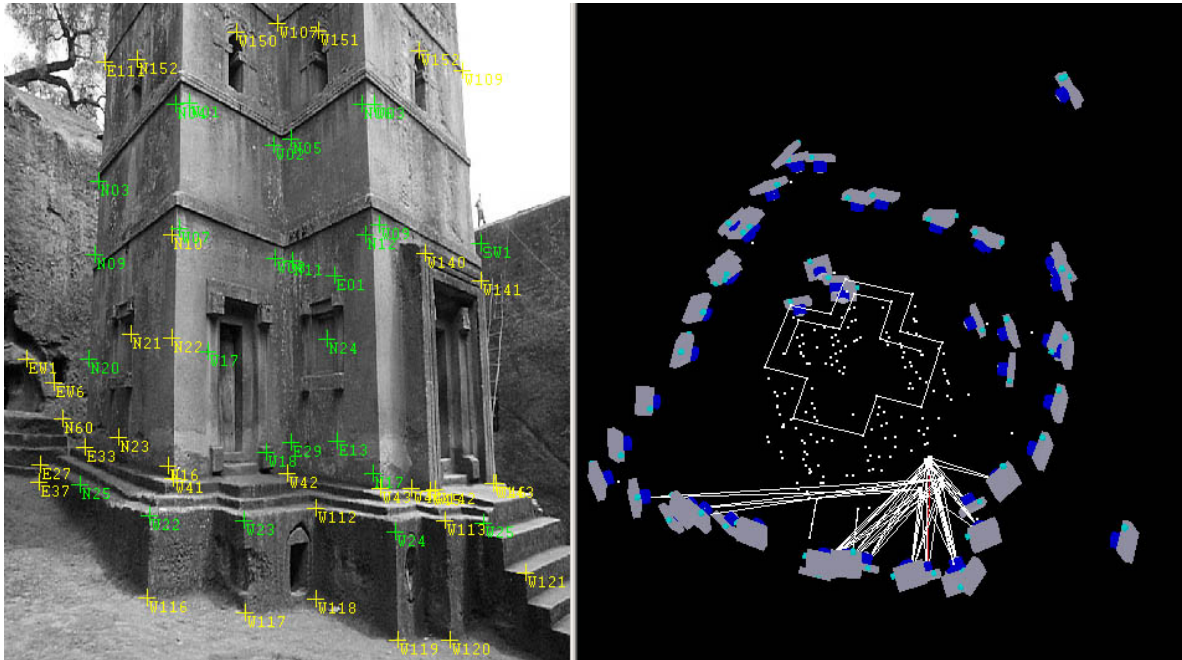


Figure 3: Multi-image convergent geometry for complex object.

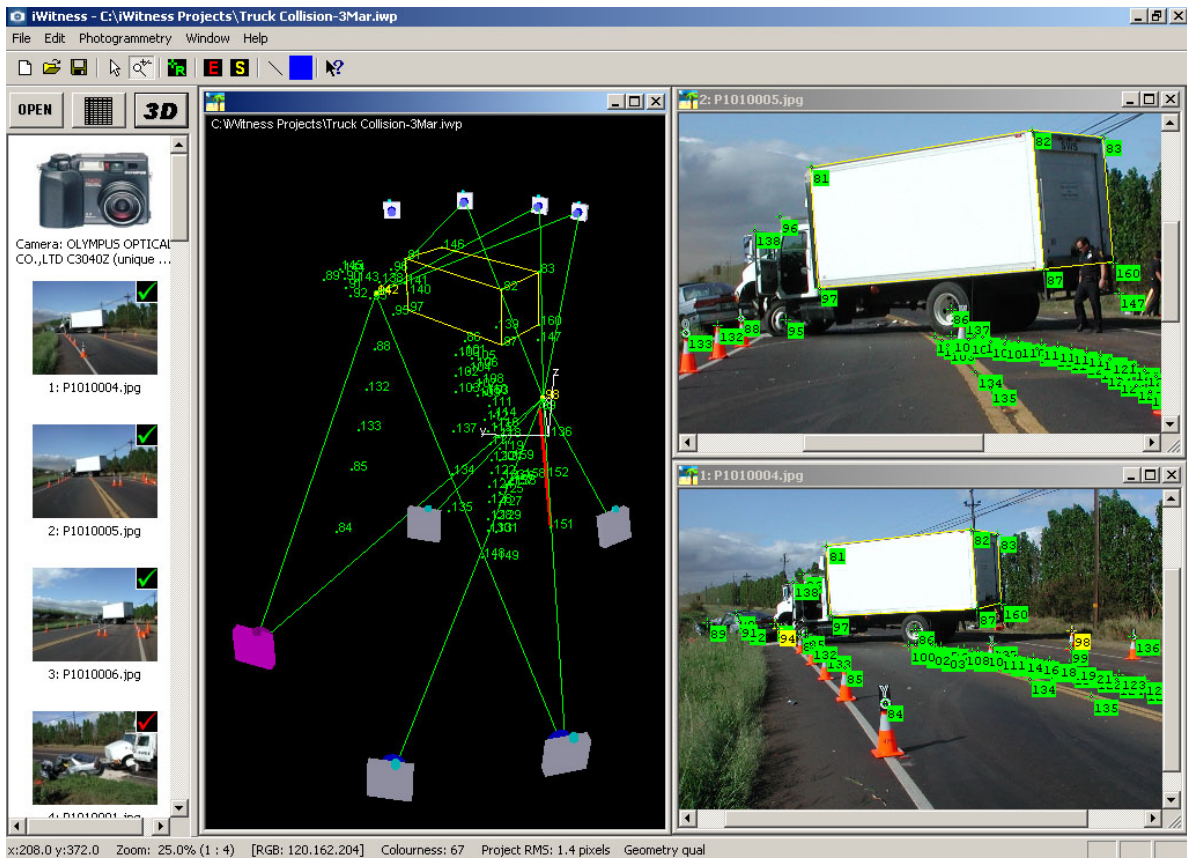


Figure 4: Multi-image geometry with camera stations near the effective plane of the object, typical in accident reconstruction.

involving closed-form solutions and initial value generation for iterative non-linear optimization. Moreover, there were different requirements relating to planar and non-planar 3D object point arrays. At the time, a fully rigorous single or multi-camera self-calibration could be obtained with much less effort and much greater model fidelity and accuracy via a photogrammetric bundle adjustment with additional parameters.

3.4 The Exterior Orientation Device and Coded Targets

The introduction of digital cameras to close-range photogrammetry opened the door to full automation of the network orientation process. Images could be scanned for targets, and after initial exterior orientation (determination of \underline{Q}) a correspondence determination based on, say, epipolar constraints or spatial intersection criteria could be employed to provide initial values for object point coordinates. But how was exterior orientation to be obtained? The answer to this question was the exterior orientation (EO) device introduced in the mid 1990s (e.g. Ganci & Hanley, 1998; Fraser, 1997). This device, examples of which are shown in Fig. 5, is an automatically detected and recognized pattern of target points within the image. The points on the EO device have known XYZ coordinates. Hence, the scenario for orientation follows that of the previous section.

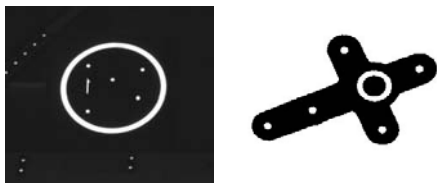


Figure 5: Examples of EO devices.

After resection of images that ‘see’ the EO device, intersection follows to determine initial values \underline{X} of object point coordinates. What happens with images in which the EO device does not appear? This is where coded targets come in. These were first proposed for close-range photogrammetry in the late 1980s. If there are coded targets distributed on the object, which are automatically recognised and triangulated in the initial spatial intersection, then groups of these become, effectively, EO devices and they facilitate resection of additional images. Once again, an iterative process of resection/intersection and possibly initial bundle adjustment is pursued until starting values for all parameters are determined. A final bundle adjustment, usually with camera self-calibration, is then performed.

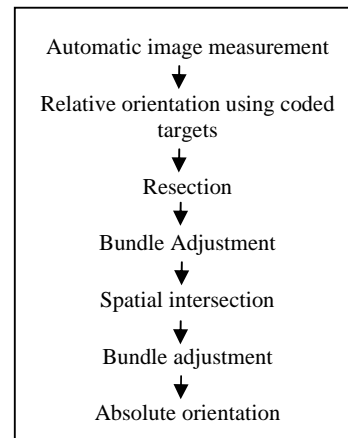
This scenario affords fully automatic network orientation and 3D coordinate determination of targeted object points. The word target here is important, for targets are essential, along with favourable illumination conditions (usually provided via a strobe flash) to ensure that the resulting image points will be automatically detected and accurately measured. Whereas from a computer vision standpoint the provision of targets is anathema, it would seem to this author that the EO device/coded target approach would go a long way to alleviating camera calibration concerns in applications such as motion tracking and object modelling, since the process is simple, fully automatic and produces an accurate, scene independent calibration.

3.5 Coded Targets Alone

In spite of the benefits of the EO device, and its widespread adoption in vision metrology, it does display shortcomings. For

very large objects, a large EO device is warranted to ensure success in the initial closed-form resection. Moreover, the camera station geometry and EO device location must be such that robust detection and measurement is provided in enough images (minimum of two) with suitable geometry to support reliable initial spatial intersection. Consider now a scenario where the EO device is not used, but coded targets are.

As with the previous process, the images would be scanned and the coded and uncoded targets measured. Because the coded targets provide point correspondences, there need only be a sufficient number of pairs of homologous points between two images to facilitate RO. This first step in the orientation process now requires a solution of the coplanarity equations (initial values $\underline{\Delta}$), with the full scenario becoming:



Following the initial RO, resection follows for those images ‘seeing’ enough of the codes included in the RO. A bundle adjustment can then be used to refine the network, after which correspondence determination and spatial intersection follow to establish additional object point coordinates (\underline{X}). This is followed by a final bundle adjustment and AO. There may be a number of resection/intersection stages in the building of the network, though these occur fully automatically and the user can be oblivious to the number of iterative cycles performed.

This process is followed in the ongoing development of the *Australis* software system, and also in *iWitness* (Photometrix, 2005; Fraser & Hanley, 2004) albeit only for sensor calibration in fully automatic form in *iWitness* (Fraser et al., 2005). While the approach has proven to be robust and reliable for providing initial values \underline{Q} and \underline{X} for the bundle adjustment, the provision of starting values $\underline{\Delta}$ for the initial RO has proven to be a very challenging problem, which has fortunately been overcome.

Given the considerable amount of literature on the essential matrix/fundamental matrix model for image orientation, one would be left with the impression that this was a viable ‘working’ approach for RO. Experience suggests otherwise, and indeed it is consistent with the observation by Horn (1990) quoted earlier. Put simply, reliable and reasonably representative values for the five parameters of relative orientation cannot be expected with a sufficient degree of confidence, especially in convergent imaging configurations with a modest number of point correspondences (say 8-10) and object arrays which display limited depth in proportion to their lateral extent. The networks in Figs. 2 to 4 provide examples of such cases.

As an alternative and remarkably straightforward approach to determining initial values $\underline{\Delta}$ for RO, a Monte Carlo type strategy has been adopted in which a very large number of possible relative orientation solutions are assessed for the available image point pairs. The refined solution in each case is obtained via the coplanarity model using combinations of plausible initial values (there could be many of these). From the number of qualifying solutions obtained for the first five point pairs considered, the most plausible are retained. But, RO results are not finalised at this time, as there may be quite a number of possible solutions in cases of weak geometry. This will be compounded by the presence of random image coordinate measurement errors, however small. The entire process takes only a fraction of a second.

3.6 On-line Orientation for Manual Image Measurement

A favourable characteristic of adopting an initial RO in the orientation process, much as is done with stereo model restitution, is that it is quite well suited to on-line initial network orientation. This is where the measurement of image coordinates involves interactively referencing points in image pairs rather than labelling them for later off-line computation of orientation. As soon as enough image points are referenced, RO can automatically take place, in the background. It is quite conceivable that in cases of very poor geometry, such as represented by Fig. 4, there may be multiple plausible solutions to the RO when only 5 to 8 points are available. The RO process must then keep track of these possible solutions and examine every one as each additional point is referenced. The correct solution is generally isolated with no more than ten points, whereas for a strong geometry a successful RO can usually be reported to the operator after 6 points are referenced. The orientation process can then be summarized as:

- 1) Mark points of interest in a single image
- 2) Reference these to a second image (no point labels necessary); RO automatically computed and updated on-line with each new point after the first 5
- 3) Reference a new image to any of the referenced images and resect after 4-6 points referenced; bundle adjustment computed and updated on-line with every new point added
- 4) Repeat (3); self-calibration can be initiated when the network is judged internally to be geometrically strong enough
- 5) Final AO to obtain XYZ coordinates in desired reference system

This on-line network orientation capability is a distinctive feature of *iWitness*. It also extends to the system's other computational processes related to orientation, namely to resection, bundle adjustment and coordinate transformation for AO. At no time need the operator select a key such as 'compute' or 'process'; this happens immediately enough information is available to support the computation. Greatly enhanced error detection is a further feature of this approach since observational blunders are recognised, and corrected, as they occur.

4. CONCLUDING REMARKS

The purpose of this paper has been to review the network orientation processes that have been employed in close-range photogrammetry as it has evolved from a film-based to digital image-based 3D measurement technology. Underlying the different orientation and sensor self-calibration algorithms and computational procedures are two basic functional models which have served photogrammetry well: the collinearity and coplanarity models. Although non-linear, both are solved via linear least-squares, thus requiring the determination of initial values for the parameters. The reason for the different computational sequences that have evolved is closely related to the different approaches to initial value determination.

It is also for this reason that one hears justification for the developments of alternative, linear solutions to image orientation. The author has attempted to demonstrate that while generation of initial values may have been viewed as an impediment in the past, it has never been much more than a necessary nuisance. Moreover, with closed-form solutions and alternative approaches to solving the relative and exterior orientation problems in an approximate manner, orientation systems requiring no operator input and no provision of additional object space information are readily realisable. For example, the provision of coded targets, not in any required configuration, is sufficient to enable fully automatic exterior orientation and camera calibration, as exemplified by the *iWitness* camera calibration process (Fraser et al., 2005).

Finally, though perhaps regrettably, this author concludes that in the context of practical close-range photogrammetry, which focuses upon metric measurement, either automatic or manual, the linear projective geometry based approaches centred upon the essential and fundamental matrices have very little to offer as viable alternative models for network orientation.

5. REFERENCES

- Abdel-Aziz, Y. & Karara, H.M., 1971. Direct linear transformation from comparator coordinates into object space coordinates in close-range photogrammetry. *ASP Symposium on Close-Range Photogrammetry*, Urbana, Illinois, 1-18.
- Faugeras, O. & Luong, Q., 2001. *The Geometry of Multiple Images*. MIT Press.
- Fischler, M.A. & Bolles, R.C., 1981. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of ACM*, 24(6): 381-395.
- Fraser, C.S., 1997. Innovations in automation for vision metrology systems. *Photogrammetric Record*, 15(90): 901-911.
- Fraser, C.S & Hanley, H.B., 2004. Developments in close-range photogrammetry for 3D modelling: the *iWitness* example, *International Workshop: Processing & Visualization using High-Resolution Imagery*, Pitsanulok, Thailand, 18-20 Nov., (on CD-ROM).
- Fraser, C.S., Hanley, H.B. & Cronk, S., 2005. Close-range photogrammetry for accident reconstruction. *Optical 3D Measurements VII, Vienna, 3-5 Oct., 9 pages*.
- Ganci, G. & Hanley, H., 1998. Automation in videogrammetry. *International Archives of Photogrammetry & Remote Sensing*, 32(5): 53-58.

- Hartley, R.I., 1992. Estimation of relative camera positions for uncalibrated cameras. *Computer Vision – ECCV'92*, LNCS Series Vol. 588, Springer, 579-587.
- Hartley, R.I., 1997. In defence of the 8-point algorithm. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 19(6).
- Hartley, R.I. & Zissermann, A., 2000. *Multiple View Geometry in Computer Vision*. Cambridge Press.
- Horn, B.K.P., 1990a. Recovering baseline and orientation from essential matrix. <http://ocw.mit.edu/OcwWeb/Electrical-Engineering-and-Computer-Science/6-801Fall-2004/Readings/>, 10 pages.
- Horn, B.K.P., 1990b. Relative orientation. *International Journal of Computer Vision*, 4: 59-78.
- Krauss, K., 2000. *Photogrammetry*. Duemmler, Vol. 1, 397 pages.
- Longuet-Higgins, H.C., 1981. A computer algorithm for reconstructing a scene from two projections, *Nature*, 293:133-135.
- Mikhail, E.M., Bethel, J. & McGlone, J.C., 2001. *Introduction to Modern Photogrammetry*. John Wiley & Sons, Inc., 479 pp.
- Photometrix: <http://www.photometrix.com.au> (Web site accessed May 20, 2005).
- Roth, G., 2004. Automatic correspondences for photogrammetric model building. *International Archives of Photogrammetry, Remote Sensing & Spatial Information Sciences*, 35(B5): 713-718.
- Stefanovic, P., 1973. Relative orientation – a new approach. *ITC Journal*, 1973-3: 417-448.
- Thompson, E.H., 1959. A rational algebraic formulation of the problem of relative orientation. *Photogrammetric Record*, 3(13): 55-59.
- Tsai, R.Y., 1987. A versatile camera calibration technique for high-accuracy machine vision metrology using off-the-shelf TV cameras and lenses. *IEEE Journal of Robotics & Automation*, RA-3(4): 323-344.