Registrering av skihopperes bevegelser ved hjelp av multiple stasjonære videokamera

Atle Nes, Ingolf Hådem[1] og Jan H. Nilsen[2]

Vitenskapelig bedømt (refereed) artikkel

Atle Nes & al.: Capturing the motion of ski jumpers using multiple stationary video cameras

KART OG PLAN, Vol. 72, pp. 56–80, P.O.B. 5003, NO-1432 Ås, ISSN 0047-3278

The paper reports on the application of photogrammetry to tracking and analysis of 3D motion of a ski jumper. Two alternative situations as to network design and computational procedure are considered: the real situation (A) and an alternative situation (B) with a simulated network design and a different computational procedure.

Situation A. Image sequences of ski jumps in the K120 large hill at Granåsen, Norway, are acquired by using a triplet of stationary, synchronized, digital video cameras. Ski jumpers have reflective markers on their jump suit, helmet and skis that serve as targets. The targets are identified and tracked in time and space to estimate dynamic parameters including position and speed of the jumper, slope of the two skis, and angle between skis and body. Accuracy is evaluated by Monte-Carlo simulation. The computational procedure resulting in estimated dynamic parameters and their accuracy, includes these steps: 1) Calibration in laboratory to estimate all interior parameters in the triplet of cameras in one single bundle adjustment, based on images of a 3D frame of unknown target points taken from different directions. 2) Pose estimation by bundle adjustment, based on a triplet of images taken by 3 properly configured cameras in an arena of 15 reference points and 6 target points (balls). 3) Intersection of ski lines and jumper points to generate a stick model for each acquired image triplet which can graphically illustrate the behavior of the jumper.

Situation B. A 2-steps computational procedure based on auto calibration, simulated network in the field and simplified assumptions on interior camera orientation is proposed and tested. The first step is a system calibration in one single bundle adjustment estimating each camera's individual pose principal distance and a non-linear distortion parameter (assuming principal point = image center). Simulated reference data are mainly limited to distances between targets on a vertical rod which is thought to being moved to given locations in the arena and exposed separately to facilitate automatic image analysis. The accuracy of dynamic parameters such as coordinates of jumper points and directions of skis is found by Monte-Carlo simulation using the same observational accuracy of image points as in the real experiment and a guessed estimate of observational accuracy of rod verticality.

Keywords: ski jumping, camera calibration, motion capturing

Atle Nes, Ph.D. student. Sør-Trøndelag University College, Faculty of Informatics and e-Learning, E. C. Dahls gt. 2, NO-7004 Trondheim. E-mail: atle.nes@hist.no

0 Introduction

Goal and overview

The goal of the present project is to test the use of multi-image photogrammetry to track and analyze 3D motion of a ski jumper on the basis of real and simulated cases in the K120 large jumping hill at Granåsen in Trondheim, Norway. The project focuses on design and application of a system to estimate, for the early flight phase, the time dependent 3D positions of strategic points and lines of a stick model of the jumper and the skis. This serves as the basis for analyzing the jumper's behavior with respect to speed, position and various angles (e.g. between the skis and the body) which may be decisive for a successful jump (see e.g. Virmavirta et al. 2005).

1. Professor emeritus. Norwegian University of Science and Technology, Faculty of Engineering Science and Technology
2. Associate professor. Sør-Trøndelag University College, Faculty of Informatics and e-Learning

Section 1 presents and discusses some relevant concepts. Sections 2–5 describe the work carried out and results obtained:

Section 2: Setup of a stationary triplet of synchronized video cameras for acquisition of a time series of image triplets in the first 20 m of flight

Section 3: Establishment of a network of (i) reference points (geodetically measured) and tie points for bundle triangulation, all points targeted in field, and (ii) recognizable markers on the jumper

Section 4: Stepwise computation starting with measured image points and surveyed field points and ending in a time series of estimated jumper points and ski lines requiring a one-sigma coordinate (X,Y,Z) accuracy $\sigma_{co} < 0.2$dm, i.e. $\sigma_{co}/d < 1{:}1000$ for a flight of d = 20 m, and

Section 5: Illustration and accuracy evaluation of the behavior parameters.

A 3-step numerical process is used. Step 1) involves camera calibration (bundle adjustment) using a 3D frame of targeted points. Step 2) involves pose estimation (bundle adjustment) of 3 cameras in the field. Step 3) involves intersection of jumper/ski features.

Section 6 reports on testing an alternative situation using a 2-steps computational procedure where the first step is a system calibration (i.e. no pre-calibration) on the basis of a simulated network in the field

Discussion, conclusions and further work are given in Section 7.

Previous work

Previous research on ski jumping shows that the most crucial parts of a ski jump are the take-off and early flight phase, corresponding to approximately 20 m in the Granåsen arena (see Fig. 1). Schwameder 2008, Schwameder et al. 2005 and Virmavirta et al. 2005 studied kinematic characteristics of the early flight phase in ski jumping. Virmavirta et al. 2005 studied the early flight phase in the large hill ski jumping competition at the Salt Lake City Olympics, 2002. They filmed a 40 m phase flight with two high-speed pan & tilt video cameras and following the jumpers for a distance of 40 m. Calibration was carried out using 2.4 m rods and a frame of 20 reference points (balls) measured with theodolite (see Peak Performance Technologies, Inc.). Calibration of image space (interior parameters) was carried out using the proper technique for pan and tilt recording: three rods placed along the area of interest, where two of these rods are aligned with the in-run tracks determining the X-axis. Accuracy, expressed as residual σ, was 0.08% (2 mm), 0.04% (1 mm) and 0.6% (14 mm) for calibration rods 1, 2 and 3, respectively. The frame was placed ca 10 m from the take-off edge, and residual σ in the x, y and z directions was 0.38, 0.00 and 0.41%, respectively. The 3-D model of the jumper's body consisted of 12 segments as is based on data from Leva 1996. The skis were digitized as additional segments. Height and speed of COM (Centre of Mass) and the 5 characteristic angles of jumper/ski geometry were analyzed and graphically illustrated from estimated segment parameters. Actual accuracies of these final data are not reported.

Hermsdorf et al. 2008 reports on a detailed biomechanical multi-body model system JUMPICUS for simulations of ski jumping. They offer various methods for biomechanical analysis of velocity, angles and trajectories. This system is based on motion capturing using video image sequences of ski jumpers with body markers. No information on video data acquisition and accuracy is presented.

A system using a carrier phase GPS receiver integrated into a helmet was developed and tested by Blumenbach 2004, 2005. It operates at 20 Hz and produces a single track showing the position of the jumper flying down the hill.

In spite of the limited research on the specific activity of ski jumping, much more research has been done in the general field of human motion analysis. Two theses (D'Apuzzo 2003 and Remondino 2006) give valuable insights into the application of photogrammetric techniques to human motion tracking. Overviews of status and advances, including concepts, procedures and techniques related to human motion analysis, are given by Wang et al. 2003 and Moeslund et al. 2006.

1 Relevant concepts

Dynamic 3D points and lines: In Fig. 6 *dynamic point* features P_1,.. P_4 and dynamic line features L1 and L2 represent a dynamic stick model of the jumper/skis and are exposed by a triplet of static video cameras (30 exposures/s) during the flight phase.

Static 3D points are those configured in the field and exposed by the camera triplet. The points act either as tie points or reference points (see below).

Tie points (Luhmann et al. 2006, p. 216): In the present project the name «tie points» is only used for unknown target points that assist the connection of two or more images in the bundle adjustment of steps 1 and 2, Fig. 4. By camera calibration (step 1) using a 3D frame, all targeted frame points act as such tie points. By pose estimation in the field (step 2), alternative 1 uses only static target points as tie points, while alternative 2 also uses positions of the dynamic helmet as tie points, see footnote 3 p. 64.

Reference points are surveyed target points (Fig. 7) referring to a geodetic network and used in pose estimation by bundle adjustment step 2, with the coordinates as random variables (conf. App. D.3).

Image processing and analysis are in the present project based mainly on visual techniques considering both point and line features. Thus the following tasks include: (i) identification of dynamic and static object *points* in 2 or 3 images of the actual triplet, determination of correspondence image ↔ object points and measurement of image coordinates by «cursor on screen» to be further used for *bundle adjustment* (BA), and (ii) identification of ski *lines* projected in 2 or 3 images of each triplet, determination of correspondence image ↔ object lines and measuring on each image line at least 2 arbitrary points, from which the parameters of the image line are derived to be further used for line intersection.

Point intersection In *(*Luhmann et al. 2006, p. 283) multi image point intersection in space is described as a least squares coordinate estimation of an object point, based on collinearity equations (Luhmann et al. 2006, section 4.2.2, p. 204), «constant» interior/exterior orientations, and measured image points that correspond over 2 or more images. The estimation minimizes the residuals in observed coordinates of image points and may thus be considered as a Bundle Adjustment (BA). This gives a non-linear problem which may be solved iteratively by the Gauss-Newton method requiring approximate, initial values for the unknown coordinates of the intersection point. However, in the present project a modified, *closed* solution as shown in Appendix C.1 is applied, i. e. no *external*, initial values are needed.

Point intersection is applied in step 3 of the flow chart in Fig. 4 to estimate points on the jumper for further calculation of the behavior parameters (Fig. 6). It is also implied in steps 1 and 2 to generate initial coordinates of tie points before bundle adjustment.

Line intersection estimates the four parameters of a 3D object line from corresponding «observed» 2D image lines each defined by two parameters . The intersection is functionally modeled by the collinearity equations for lines. These equations are derived in (Luhmann et al, 2006, chapter 4.5.2, p. 296) on the basis of image and object line equations and collinear equations for points. To avoid ill-conditioning the line equations must be formulated such that the direction parameters are small. With the object system defined as Fig. 8 the formulae of Luhmann et al, 2006 do not fulfill this requirement of the ski line equations and must therefore be modified to Eq. (C9) as shown in App C.2.

The direction parameters of a ski line determine the ski slope (Fig. 6) and contribute also to determine sufficient accurately the ski line segment (graphically illustrated in Fig. 8) from i) the intersected point 3 or 4 (Fig. 6) near the ski binding ≈ midpoint of the ski and ii) the ski length = 2.63 m. The least squares solution minimizes the coordinate residuals in observed line points (see App. C.2).

Line intersection is executed in 2-steps: Step A estimates the 2D line parameters by a closed, iterative numerical solution separately for each line on the basis of the line Eq. (C7) and observed image line points, and step B estimates the 3D line parameters by an iterative numerical solution on the basis of non-linear collinear Eq. (C9) or (C10), approximate values of unknown line parameters and the output of step A (see App C.2). A least squares solution

that minimizes the coordinate residuals in observed line points is applied. In the present project this procedure has been programmed and successfully applied. A one-step solution procedure is mentioned in Notes, App. C.2.

In addition, an alternative estimation of initial values by line intersection in homogenous presentation (Hartley and Zisserman 2003, section 12.7, p. 321) has been considered and tested: Defining the object line as the axis of a pencil of planes, each of which constitutes the joint of the corresponding image line and projection center, good initial values of the line parameters are found by intersecting those two planes that Singular Value Decomposition proves to give the «best» intersection result. The formulation is shown in App. C.3. In the present project a procedure for checking the line correspondence over three images as described in (Hartley and Zisserman 2003, pp. 365–367) have been programmed and successfully tested. By this procedure, interchange of left/right ski due to bad image quality has been revealed.

Line intersection from three views minimizing the distances in image between measured and projected estimated lines as described in (Hartley and Zisserman 2003, p. 323) has not been considered in the present project.

Space resection with unknown interior orientation (also called DLT by Abdel-Aziz & Karara 1971 and Luhmann et al. 2004, p. 212) is used for approximate pose estimation of single images before bundle adjustment steps 1 and 2 of Fig. 4. Formulating the collinearity equations as homogeneous linear equations in 12 parameters F_i which are the elements of 3x4 projection matrix $\boldsymbol{P}$ (see Eq. (A2b) App. A), the method of deriving approximate pose parameters is:

– In the first place solve linearly 11 parameters (e.g. $F'_i = F_i / F_{12}$ (i = 1..11) assuming $|F_{12}| >> 0$) by a least squares fit based on at least 6 non-coplanar, known object points (X, Y, Z) and corresponding measured image points (x, y).
– Secondly, from estimated F'_i derive approximate pose parameters. To avoid instability due to $|F_{12}| \approx 0$, the homogeneous solution method that estimates F_i (i = 1..12) up to unknown scale (asprs 2004, p. 778) is preferred.

Space resection with known interior orientation (asprs 2004, p. 786 and Shigang Liu et al. 2008) is used as a *direct* (non-iterative) procedure for pose estimation of a single image from: (i) n ≥ 3 known, non-collinear points (X,Y,Z), (ii) measured image coordinates of corresponding image points (x, y) and (iii) known interior orientation. When n = 3 ambiguity may appear, so n > 3 is advisable. In the present project this resection approach was successfully tried as an option to the DLT-method, to derive initial poses before bundle adjustment, step 2.

Relative Orientation (RO). Luhmann et al. 2006 p. 218 *describes RO as «the translation and rotation of one image with respect to the stereo-partner in a common local model coordinate system».*

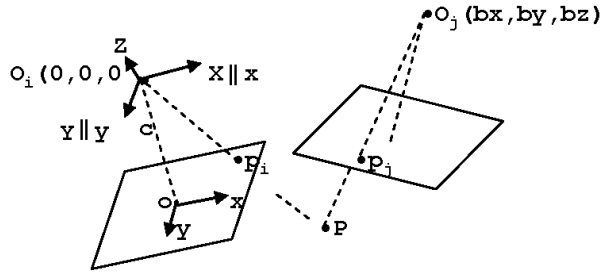


Figure 1 Geometry of stereo model. Projection centres (O). Homologue image points (p). Model point (P). Principal distance (c), Principal point (o). Image system (x,y). Model system (X,Y,Z). Basis components (bx,by,bz)

RO is used to start the derivation of initial values before bundle adjustment (BA) in step 1

(camera calibration) by finding such values for the 12 outer orientation parameters ($\mathbf{R}$, $\mathbf{X}_0$), ($\mathbf{R}$, $\mathbf{X}_0$)$_j$ of two selected bundles (i, j) in a stereo model system (XYZ), see Fig.1 where the left projection center is chosen as origin and $\mathbf{R}_i = \mathbf{I}$. With

$$(\mathbf{X}_0)_j = (\mathbf{X}_0)_i + \mathbf{b},\ \mathbf{b} = (bx,by,bz)^T$$

the algorithm is:

- choose fixed values of ($\mathbf{R}$, $\mathbf{X}_0$)$_i$ e.g. ($\mathbf{I}$, $\mathbf{0}$)
- choose a fixed value, (e.g. '1 ') of bx (if bx ≈ 0, interchange X and Y)
- estimate iteratively the remaining 12–7 = 5 parameters ($\mathbf{R}$, Y_0, Z_0)$_j$ on the basis of
  - *i) measured n ≥ 5 pairs of homologue image points*
  - *ii) coplanarity condition (Luhman et al 2006, Eq (4.32) p. 219).*
  - *iii) the minimization constraint on residuals ε_x, ε_y in image coordinates x, y*
    $\Sigma(\varepsilon_x^2 + \varepsilon_y^2)$ min. Σ over measured image points
  - *iv) approximate values of ($\boldsymbol{R}$, Y_0, Z_0)$_j$*

The present project practices a *closed-form* iterative solution. It tries and fails a set of different, probable $\mathbf{R}_j$ together with linearly derived (Y_0, Z_0)$_j$ as initial values (see Nørbech 1982). When n = 5 ambiguity may occur. When n > 5 it may succeed in a unique, iterative least squares solution minimizing the square sum of residuals in image coordinates, thus facilitating accuracy and reliability analysis (Luhmann et al. 2006, p. 64). Hådem 1984 and Horn 1987, 1990 and 1991 report on *continuation methods* for optimal solution of the relative-orientation problem without requiring a *good* initial estimate. Such methods were not used in the present project.

Bundle adjustment (BA), (Hådem 1989, Luhmann et al. 2006, pp. 21, 52, 229) BA is generally considered as a *one-step* least squares simultaneous estimation of all unknowns by taking optimally into account (i) all functional relations between parameters, and (ii) stochastic properties of observations (see App D.1). BA is *functionally* based on (i) the collinear equations with parameters like elements of interior/exterior orientations and coordinates of image/object points, and (ii) additional equations, e.g. functional relation between a distance and coordinates of its end points. The total equation system is considered as the functional part of the mathematical model. Observational data of e.g. image and object coordinates is considered as sample of stochastic variables with the corresponding parameters in the functional model as expectations. The stochastic part of the mathematical model is expressed as a variance / covariance matrix of the stochastic variables. BA implies a non-linear solution problem which requires linearized equations and approximate values of all unknowns to initialize the iterative solution process.

Free bundle adjustment is characterized as BA with insufficient external-given reference data to define the 7 parameters of the object system. In (Luhmann et al. 2006, p. 248) it is shown how to eliminate the rank defect by extending the normal equations by 7 lines expressing a differential similarity transformation and minimizing the translation corrections of new points. This measure does not change the internal quality of the block as to image coordinate measurement and interior orientation.

For free BA applied in the present project the *Weighted Constraints* approach (asprs, p. 881) is implemented: The initial, approximate values of both outer orientation parameters and unknown point coordinates are considered as a sample of independent «observations» with guessed accuracy and formulated as weighted constraints . (Formulation of such constraints for outer orientation, see App D.2). The addition of such constraints may eliminate the rank defect that is due to external, insufficient reference data and may also generally improve the numerical condition.

In step 2 (pose estimation by BA, see Fig 3) the surveyed coordinates of static target points are taken into account through additional constraints ((D3b), App D) to define the 7 parame-

ters of the object system in which the jumper points are finally estimated by intersection in step 3. App D.3 formulates optimal estimation of the stochastic weight relation between measured image coordinates and surveyed object coordinates. These relations are taken into account in the BA-estimation of the poses.

Free BA is applied in step 2 as a means of *gross error search* (Luhmann et al 2006 sect 2.3.4 p. 67) as follows:

- Free BA (ignoring the field survey) may reveal observational gross errors in image coordinates.
- Then, 3D transformation of the free bundle adjusted block of coordinates and poses on surveyed reference points may reveal gross errors in reference coordinates.
- Final BA with additional constraints of reference data (and thereby improved redundancy and geometry) may reveal further gross errors in 2D – and 3D coordinates.

Stepwise estimation of the dynamic points and line features of jumper/skis is applied as follows.

In the real experiment as a 3-steps procedure (see flow chart, Fig. 4):

Step 1) Free BA (with no extern known reference data) to estimate camera-dependent interior orientations by the use of a calibration frame as a test field of target points which act as (unknown) tie points:

Step 2) Estimate by BA (with full extern reference data) the poses in the jumping field, with the result from step 1 as «error free»

Step 3) Estimate by intersection the jumper points and ski lines with the result from steps 1 and 2 as «error free»

In section 6 '*Simulated network and system calibration'* a two-step procedure is applied:

Step1' Estimate by BA the interior/exterior orientations on the basis of (i) points like P_0, P_1 ,P_2 and jumper helmet (Fig. 9) acting as tie points and (ii) additional constraints like (3) for the verticality of line $P_1 - P_2$ and measurement of distance $P_1 - P_2$.

$$\left.\begin{aligned} E(\Delta X_{12}) &= (X_1 - X_2)_{unkn} \\ E(\Delta Y_{12}) &= (Y_1 - Y_2)_{unkn} \\ E(d_{12}) &= Abs(Z_1 - Z_2)_{unkn} \end{aligned}\right\} \quad (3)$$

where the expectations E(.) of stochastic coordinate components ΔX_{12}, ΔY_{12} and distance d_{12} are functions in unknown coordinates of P_1, P_2.

Step2' Estimate by intersection the jumper points and ski lines, with the result from the step1' as «error free».

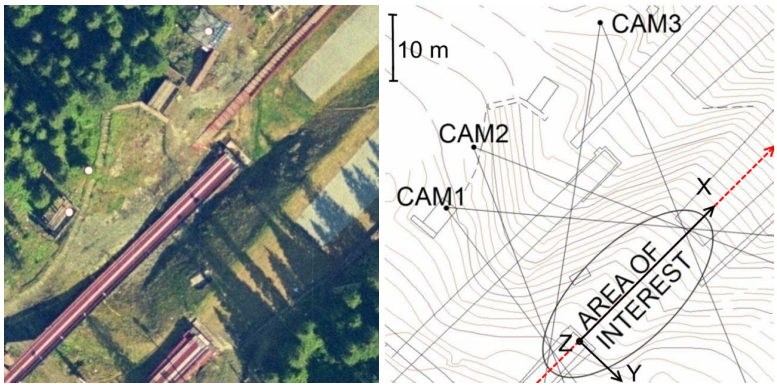
2 Setup for a stationary triplet of synchronized video cameras

Equipment

Our current setup consists of three identical IIDC DCAM compatible FireWire IEEE 1394a Allied Vision Technologies Marlin F-080B video cameras, each delivering 8-bit grayscale video with an image size of 640 x 480 pixels (VGA) at 30 fps. Direct digital output eliminates the need for a dedicated frame grabber card. The cameras are connected to a single computer through three sets of Optics M4-200 optical FireWire IEEE 1394b repeaters and 400 m multimode optical fibers. This enables great flexibility in camera positioning and solves the problem of FireWire cable length limitation (max 10 m) for each camera on the hill.

Camera configuration

Optimal camera configuration is described in Olaugue 2002 and Frazer 1996. In our case, the cameras are stationary. They are positioned and oriented in the jumping hill (Fig. 1) such that they point toward the scene circumscribing the early flight motion of the ski jumper. Once properly adjusted, no camera rotation, focusing or zooming is allowed. Notice that all cameras are placed on the same side to ensure that the jumper point features to be 3D estimated are «seen» by at least two cameras. It is generally advisable to use three or more cameras for triangulation, to ensure reliability of the image point measurements.



(a) (b)

Figure 2 (a) Aerial photo showing the K90 normal ski jumping hill to the left and K120 large hill to the right. Map to the right (b) shows the camera positions, their fields of view and the coordinate axes.

Synchronization

To achieve high accuracy of 3D measurements of moving feature points, it is important that corresponding image frames captured from different camera views are captured synchronously. This is especially important if the objects are fast-moving, like ski jumpers, and features can move across many image pixels in just a short time period. Notice, on the contrary, that reconstruction of static 3D feature points requires no synchronization at all.

Two shielded RG58 coaxial cables (200 m) link the cameras for hardware synchronization. This allows an external triggering pulse to travel to all the cameras, effectively synchronizing the three video streams. The captured video images, originating from different views at a specific time, are concatenated sideways into one single frame and tagged to simplify later identification (Fig. 3).

Camera 1 Camera 2 Camera 3

Figure 3 An image triplet used for 3D reconstruction of a ski jump

3 Network design

Establishment of a static network of signalized points in the jump arena

Some of these points serve as references, others as tie points for pose estimation by photo triangulation. All points were signaled by placing reflective-coated handballs at the specified locations. These can be seen as small white dots in the video images (Fig. 3). Reference points in the hill were surveyed with high accuracy using two real-time kinematic stations and a robotic total station. Signaled features should be spherically shaped to most accurately calculate their centre of gravity.

Establishment of dynamic target points on the jumper

Our experiments showed that reflective markers were the easiest to distinguish on dark ski jump suits. Strategically positioned reflective stickers attached to different parts of the jump suit serve as target points that can be followed in time and space.

4 Algorithm and flow chart

The algorithm consists of 3 main steps (see flow chart in Fig. 4):

Step 1 Camera calibration (Luhmann et al. 2006, p. 448) with a 3D frame (Fig. 5) as test field; the computation consisting of the following operations to obtain an optimal estimation:

1A Compute (introduce) approximate interior orientation (Luhmann et al. 2006, p. 114):
- a) Compute principal distance as in App. B, with test field Fig B2 like any side of frame Fig 4.
- b) Introduce principal point ≈ image center and affinity ≈ shear ≈ lens distortion ≈ 0.

1B Compute approximate poses (Luhmann et al. 2006, p. 202) with interior orientation known from operation 1A.
- a) Compute two camera poses of a starting stereo-model in local model system by *relative orientation*
- b) Compute coordinates of remaining unknown target points by *intersection* from known poses
- c) Compute remaining unknown camera poses by *space resection* from known target points

If a target point or a camera pose is still unknown, go to step 1Bb) or 1Bc)

1C Do free *bundle adjustment (*BA*) (App D.2)* with
- a) measured image coordinate as equally weighted observations
- b) interior orientations (camera dependent), and outer orientations (image dependent) as orientation unknowns
- c) 3D coordinates of frame points as coordinate unknowns
- d) initial, approximate values of all unknowns from operation 1A

Step 2 Pose estimation in field (Fig. 7); the computation consisting of the following operations:

2A Compute approximate values before BA
- a) Compute camera poses by resection from *reference points*
- b) Compute position of tie points[3)] (p 64) by intersection with known poses obtained in operation 2Aa

2B Do BA with additional constrains of surveyed coordinates of reference points and with:
- a) measured image coordinates equally weighted and surveyed reference coordinates with derived weights (conf. App D.3)
- b) interior orientations as «constants» and outer orientations (poses) as orientation unknowns

c) 3D coordinates of target points (tie points[3)] and reference points) as coordinate unknowns
d) initial, approximate values of all unknowns from operation 2A

Step 3 Compute locations of jumper points and ski lines by separate intersections (App C)

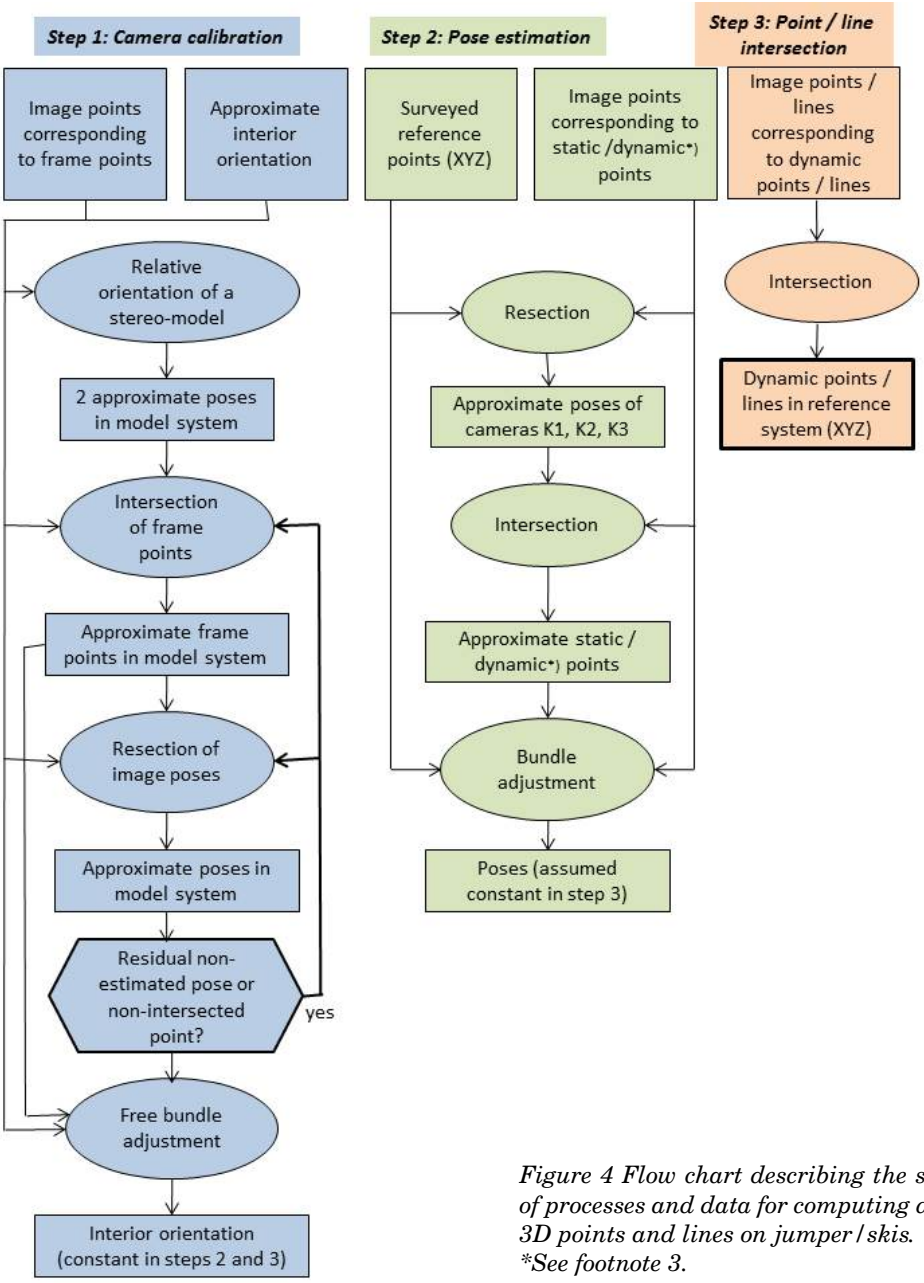


*Figure 4 Flow chart describing the sequence of processes and data for computing dynamic 3D points and lines on jumper / skis. *See footnote 3.*

3. The poses are computed in two alternatives. In alternative 1 only static target points act as tie points. In alternative 2 the dynamic helmet locations are also considered as tie points.

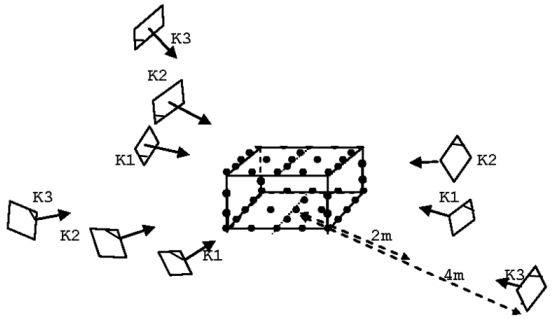


Figure 5 Sketch of network for camera calibration. 9 images taken by 3 cameras: K3 at distance 4m and K1 and K2 at distance 2m from a 3D frame (0.5m x 0.5m x 0.3m). Target points (•). Rotations of cameras are indicated.

Stick model of jumper/skis

Geometry and definition of geometrical parameters of the ski jumper and skis are shown in Fig. 6.

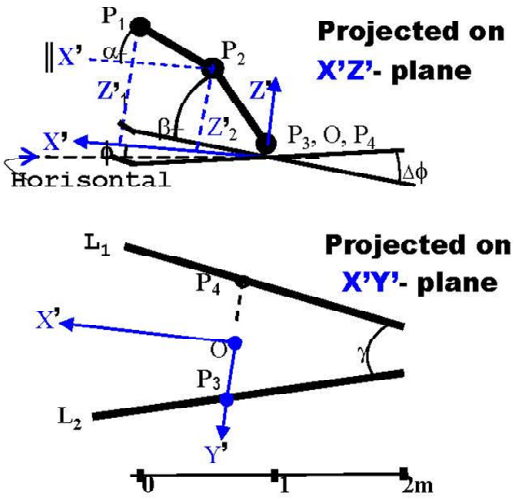


Figure 6 Geometry of the stick model of jumper / skis

Points P_1, .. P_4 and ski lines L_1 and L_2 are estimated in static system (XYZ), Fig. 1b with horizontal XY-plane. The 6-parameter dynamic system $(X',Y',Z')^T = \mathbf{R}_{\phi\omega\kappa}(X - X_0, Y - Y_0, Z - Z_0)^T$ has origin O at the midpoint of line $P_3 - P_4$, X'- axis direction = mean of the directions of lines L_1 and L_2, and $Y'(P_1) = 0$. Parameters of primary interest are 1) $\alpha, \beta, \gamma, \Delta\phi$: angles between model components of jumper/skis, 2) ϕ: slope of X'-axis, 3) Z_0: height of dynamic point O and 4) v: speed of this point.

5 Presentation of result and accuracy evaluation

Tracking parameters and their accuracy

Fig. 7 shows the basic network, and Fig. 8 shows a graphical presentation of some numerical values of the tracking parameters and the accuracy of estimation of these parameters.

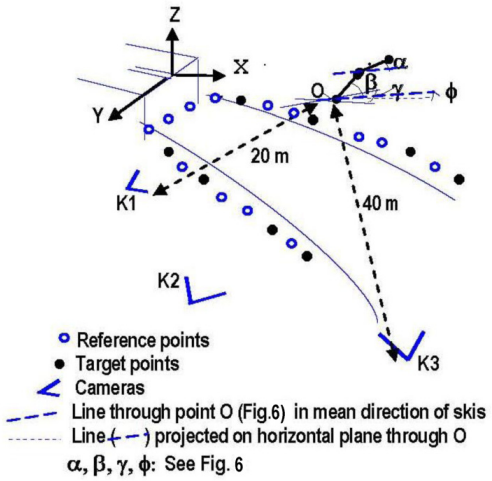


Figure 7 Sketch of arena with network for estimation of poses and reconstruction of a stick model of jumper / skis from an image triplet

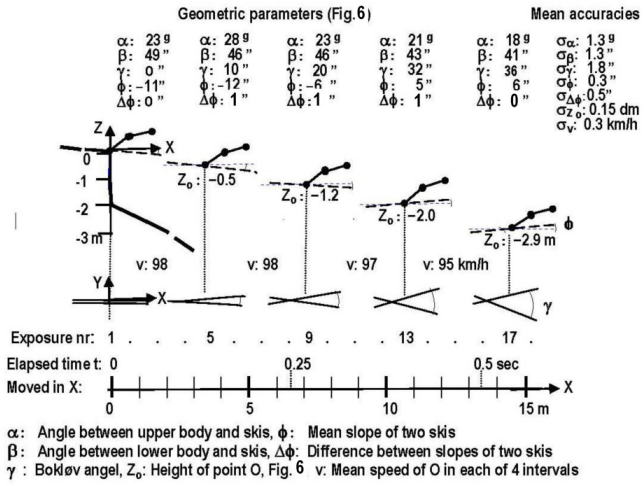


Figure 8 Stick model projected on the XZ- and XY-planes at 5 evenly distributed exposures. Speed v based on 30 exposures per second.

Observational accuracy

Accuracy σ_{xy} in image coordinates, step 1: From residuals $\mathbf{e}_{xy}$ in image coordinates σ_{xy} is derived as

$$\sigma_{xy} = \sqrt{(\mathbf{e}_{xy}{}^{T}\mathbf{e}_{xy}/r)}$$

where r = (2n – u) is the redundancy and $\mathbf{e}_{xy} = ((e_x, e_y)_1, .. (e_x, e_y)_n)^T$ is a vector of 2n coordinate residuals in n image points after free BA (conf. App D.2)

u = the number of unknown frame point coordinates + the number of interior/exterior orientation parameters – 7 internal reference constraints.

Result: $\sigma_{xy} = 0.18$ pix

Accuracy σ_{xy} in image coordinates and σ_{XYZ} in reference coordinates, step 2 is found as (see App. D.3):

$\sigma_{xy} = 0.33$ pix, $\sigma_{XYZ} = 0.12$ dm

Accuracy of interior/exterior orientations
Some figures are given in Table 1.

	Principal distance		Principal point		Camera position		
Camera	c (pix)	σ_c	σ_{x_0}(pix)	σ_{y_0}	σ_{X_0}(dm)	σ_{Y_0}	σ_{Z_0}
K1	1419.7	1.9	3.8	3.8	0.43	0.48	0.42
K2	1421.7	2.0	3.8	3.7	0.44	0.52	0.39
K3	2493.3	5.	10.	11.	0.84	0.69	0.72

Table 1 Accuracy (σ) of interior parameters c, (x_0, y_0) derived in the lab and camera positions X_0, Y_0, Z_0 derived in the field

Accuracy of jumper positions
Some figures for three jumper positions are given in Table 2.

Alter-	Jumper position 1			Jumper position 9			Jumper position 17		
native*)	σ_X	σ_Y	σ_Z	σ_X	σ_Y	σ_Z	σ_X	σ_Y	σ_Z
1	0.25	0.20	0.15	0.15	0.15	0.15	0.25	0.25	0.20
2	0.15	0.15	0.10	0.10	0.10	0.10	0.20	0.20	0.15

*Table 2 Mean coordinate accuracy (σ in dm) of jumper points P1, .. P4 (Fig. 6) for three jumper positions at time ti , i = 1, 9, 17 (Fig. 8). *) See footnote 3 p 64*

Note: The accuracy of camera positions (Table 1) and jumper points (Table 2) are derived by the Monte-Carlo method (Hastedt 2004) assuming «constant» interior orientations, observational coordinate accuracy $\sigma_{xy} = 0.33$ pixels in image points and $\sigma_{XYZ} = 0.12$ dm in reference points (see App. D.3*)*.

6 Simulated network and system calibration

For simulation of bundle adjustment, see (Luhmann et al. 2006, p. 264, section 4.3.6.1).

An alternative procedure based on simultaneous estimation of interior and exterior orientation (system calibration) and simulated network in the field is considered.

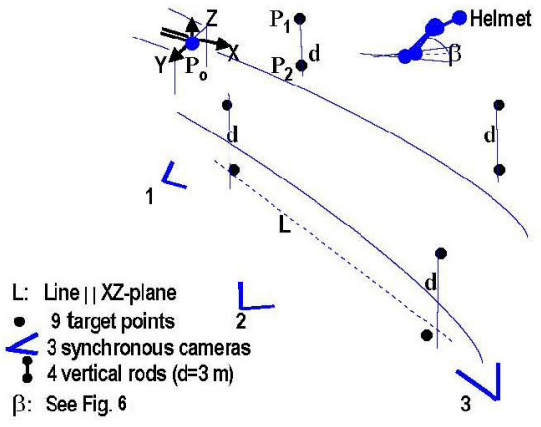


Figure 9 Sketch of arena with network for system calibration and reconstruction of a jumper/ski model from an image triplet.

Particular assumptions
- A network is simulated as in Fig. 9 with object system (XYZ) defined by (i) point P_o below the take-off edge (chosen as origin), (ii) 2 rods aligned parallel to XZ-plane with the longitude in-track direction as the X-direction and (iii) vertical rod direction as the Z-direction.
- 2 unknown interior parameters per camera: principal distance and radial lens distortion (Luhmann et al. 2006, p. 114) while principal point = known image center, and affinity = share = 0
- System calibration using bundle adjustment (BA) with constraints (3), section 1 and with helmet points as additional tie points, see footnote 3.
- Observational accuracy:

line verticality:	$\sigma(\Delta X_{12})/d = \sigma(\Delta Y_{12})/d = 1{:}100\ (0.6^g)$
distance d:	$\sigma_{\mathbf{d}} = 0.01$ dm (nearly error free)
image coordinates:	$\sigma_x = \sigma_y = 0.33$ pix (as in the real case)

Under these particular assumptions the Monte Carlo derived estimation accuracy met the required accuracy (as specified in 'Introduction'): a one-sigma coordinate (X,Y,Z) accuracy $\sigma_{co} <$ 0.2 dm, i.e. $\sigma_{co}/d < 1{:}1000$ for a flight of d = 20 m .

Proposed image acquisition and image measurement for system calibration in field
We assume that a movable rod is successively placed into the 4 positions, Fig. 8. A separate image triplet is acquired for each position, to identify the two retro reflex targets, determine correspondence image ↔ object and perform digital image measurement (semi-) automatically on the basis of recognizable features. The four separate triplets represent thus one triplet of all rod positions.

Proposed derivation of initial values for system calibration:
- *Principal distance* may be derived as in App. B
- *Camera poses* may be derived stepwise: (i) relative orientation (*RO)* of a stereo-model, say 1/2, (ii) intersection of model points, (iii) absolute orientation on vertical distances (d), line L || XZ-plane and point P_o chosen as origin (Fig.8), and (iv) resection of image 3.

Note
The minimum network required depends generally on chosen behavior parameters (Fig. 6). Thus,
- angles α, β, γ, $\Delta\varphi$ needs no data
- ski slope φ needs data for obtaining a horizontal XY-plane
- velocity v of jumper point O (Fig. 6) needs known object distance to get proper scale of moved distance of point O.

7 Discussion, conclusion and further work

Summary of tasks and computation procedures
The present project dealt with theoretical and practical aspects of using a triplet of video cameras for capturing and analyzing the behavior of a ski jumper in a flight of 20 m in Granåsen jump arena. Two alternative situations were tested. The main tasks in A were: i) support of a 3D calibration frame in the lab, ii) establishment of a network of static reference points and tie points, iii) image acquisition and measurement, followed by a 3-steps computation: 1) camera calibration in the lab, 2) pose estimation of the camera triplet in the field and 3) intersection of jumper points/skis to derive the behavior parameters. Situation B differed from A by merging steps 1 and 2 into one step as system calibration based on a simulated network in the field.

Comparing situations A and B; conclusions
The main advantages of situation B compared with A are:

- It saves the cost of pre-calibration in the lab. (In the present project it involved about 1/3 of the total cost).

– System calibration reflects better the real conditions in the field (e.g. distance camera – object points) and may take into account more easily focusing and zooming.

On the other hand situation B requires a sufficiently strong geometry of the network taking into account that both interior and exterior orientations are unknowns. With good cameras it may be justified to reduce the number of unknown image errors (e.g. consider only one parameter for non-linear distortion) and also use dynamic points on jumper as tie points, to avoid possible over-parameterization and to get target points in most of the central parts of the measuring volume.

Separately acquired «calibration images» and «jumper images» require stabile orientations over time. Establishing a permanent network for updating the calibration during the jumping would be helpful, but requires efforts in designing, positioning and maintaining the corresponding target point network in the arena.

Difficulties met
Jumper points and line features were sometimes hardly identifiable in all images of a triplet. The reason may be a) the small image scale (adapted to fully cover the first 20 m of the flight), b) unfavorable view angles, c) non-optimal conditions of shadows and sunlight and d) occlusion. One particular problem was to know if an identified ski was the left or the right one.

Software program development
A photogrammetric software package FOMACON, containing both constrained and free bundle adjustment, Hådem 1989, was used together with a developed software, programmed in java and Matlab scripts.

Accuracy, conclusions
The project achieved an *observational* σ*-accuracy* of *0.2–0.3 pixel* in coordinates of image points when these points correspond to retro reflex targeted frame points (camera calibration, step 1), ball targets in field (pose estimation, step 2) and specific recognizable marks on the jumpers dress and helmet (point intersection, step 3). This result is comparable to the observational accuracy obtained in the literature of close range photogrammetry based on targeted triangulation points (Luhmann et al. 2006, p. 442).

The following result was obtained for *positional accuracy of jumper points* by *alternative 2* (see footnote p. 64).

Mean accuracy of 5 jumper heights Z (Fig. 8): $\bar{\sigma}_z$ = *0.15 dm*

Mean relative accuracy over distance d = 20 m: $\bar{\sigma}_d$/d = *1:1300*

with $\bar{\sigma}_d = \sqrt{((\bar{\sigma}_X^2 + \bar{\sigma}_Y^2 + \bar{\sigma}_Z^2)/3)}$: Mean accuracy of dynamic jumper point coordinates, based on Table 2.

We conclude that the *accuracy* result meets the one-σ requirement of 0.2 dm for jumper height accuracy and 1:1000 for relative accuracy for a distance of 20m (see Introduction). *Alternative 1* gave ca 30 % lower accuracy than alternative 2 (see Table 2).

Future work
Future improvements of the system might include improved time and space resolution, a greater number of cameras on both sides of the jumping hill and based on automatic multi-camera system calibration (Hartley & Zisserman 2003 pp. 458 and Maas 1998). A higher degree of automatic image processing and analysis is desired, e.g. the application of Hough transformation technique (e.g. Fernandes and Oliveira 2008, Sæther et al. 1991) to recognize object lines, or the technique of subtracting the image of pure static features from the image of complete static and dynamic features (e.g Piccardi 2004 and Carr 2008). Further research is desired on improved targeting of points and lines in the field, automatic recognition and digital measurement. Physical marketing on jumper/skis should be abundant, and rather use marker-less motion capture techniques as described in (e.g. Moeslund et al. 2006, Poppe 2007).

8 Acknowledgements
The authors would like to thank Beatrix Vereijken, Steinar Bråten and Gertjan Ettema at Program for Human Movement Science at NTNU for valuable insights into the art of ski jumping. Also thanks to the Norwegian national ski jump team and Mika Kojonkoski for allowing us to accompany the outdoor training and capture their ski jumps. All data presented in this article were acquired in co-operation with Kristian Brenden and ski jumpers from Heimdal high school. The calibration frame for calibrating the cameras was borrowed from SINTEF Fisheries and Aquaculture. It has been used to study ocean currents in a related project called Hydriv (Løvås 2003, Nilsen & Hådem 1994). We would also like to thank Prof. McClimans and Prof. Skramstad at NTNU for valuable comments.

9 References

American Society of Photogrammetry and Remote Sensing (asprs) *Manual of Photogrammetry* (ed. Mc Glone) 2004, asprs USA.

Abdel-Aziz, Y. I. & Karara, H. M. (1971). *Direct linear transformation from comparator coordinates into object space coordinates in close-range photogrammetry.* ASP Symposium on Close Range Photogrammetry, pp. 1–18

Blumenbach, T. (2004) *High precision kinematic GPS positioning of ski jumpers*. ION GNSS 2004, pp. 761–765

Blumenbach, T. (2005) *GPS-anwendungen in der sportwissenschaft. Entwicklung eines messverfahrens in das skispringen.* Technischen Universität Dresden, München

Carr, P., (2008) *GPU Accelerated multimodal Background Subtraction*, in Proceedings of the International Conference in Digital Image Computing: Techniques and Applications, 2008, pp. 279–286.

D'Apuzzo, N. (2003) *Surface measurement and tracking of human body parts from multi station video sequences.* Swiss Federal Institute of Technology, Zurich, Switzerland.

Frazer, C.S. (1996) *Network Design*. In: Atkinson (ed.) *Close-range Photogrammetry and Machine Vision*. Wittles Publishing, UK, pp. 256–282.

Fernandes, L. A. F. and Oliveira, M. M., 2008. *Real-Time line detection through an improved Hough transform voting scheme.* Pattern Recognition Vol. 41, pp. 299–314.

Hartley, R. and Zisserman, A. (2003) *Multiple View Geometry in Computer Vision*. Second edition, Cambridge, Univ. Press, ISBN 0521 54051 8

Hastedt, H. (2004) *Monte-Carlo simulation in close range photogrammetry.* International Archives of Photogrammetry, Remote Sensing and Spatial Information Science, 35(B5), pp. 18–23.

Hermsdorf, H, Hildebrand, F, Hofmann, N and Muller, S. (2008). *The Engineering of Sport 7.* Springer Paris. ISBN 978-2-287-09411-8, pp. 491–497

Horn, B. K. P. (1987) *Relative Orientation*. Memo 994, Artificial Intelligence Laboratory, MIT, Cambridge, Massachusetts, November.

Horn, B.K.P. (1990) *Relative Orientation*, International Journal of Computer Vision, Vol. 4, No. 1, p. 59

Horn, B. K. P. (1991) *Relative orientation revisited*. In: Journal of the Optical Society of America, A, October 1991. Vol. 8, pp. 1630–1638

Hådem, I. (1984) *Generalized Relative Orientation in Close Range Photogrammetry.* Presented paper, ISPRS-Congress. Commmision 5, Rio-Brasil. In: Int. Arch. Photogramm. and Remote Sensing, XXV(A5), pp. 372–381.

Hådem, I. (1989) *FOMAKON, a program package for industrial photogrammetry.* Kart og plan, 49(1), pp. 25–34.

Koch, K.R. (1999) *Parameter Estimation and Hypothesis Testing in Linear Models*. Second Updated and Enlarged Edition. Springer.

Leva, P. De. (1996) *Adjustments to Zatsiorsky–Seluyanov's segment inertia parameters.* J. Biomech. 29, pp. 1223–1230

Luhmann, T., Robson, S., Kyle. S and Harley, I. (2006) *Close Range Photogrammetry.* Whittles Publishing. ISBN 1-870325-50-8

Løvås, S. M. (2003) *HYDRIV task 4: modernization and improvement of 3D particle tracking using three synchronous cameras for near real-time analysis*. SINTEF Fisheries and Aquaculture, Coastal and Ocean Engineering.

Maas, H,G, (1998) *Image sequence based automatic multi-camera system calibration techniques.* Int. Archives of Photogrammetry and Remote Sensing and spatial Information sciences. Vol. XXXII, Part 5, pp. 763–768.

Moeslund, T. B., Hilton, A. and Kruger, V., (2006). Computer Vision and image Understanding. Volume 104, 2–3, pp. 90–126.

Nørbeck , T. (1982) *Generell relativ orientering* (In Norwegian). Thesis, Norwegian University of Science and Technology, Trondheim

Nilsen, J. H. & Hådem, I. (1994) *Photogrammetric tracking of tracer particles in modelled ocean flows*. ISPRS Journal of Photogrammetry and Remote Sensing, 49(6), pp. 9–20.

Moeslund, T. B., Hilton, A. & Krüger, V.(2006) *A Survey of Advances in Vision-Based Human Motion Capture and Analysis.* In: *Computer Vision and Image Understanding*. 104, 2–3, p. 90–127.

Olague, G. and Mohr, R. (2002) *Optimal Camera Placement for Accurate Reconstruction*. Pattern Recognition, Vol. 35 (4) (2002) pp. 927–944.

Peak Performance Technologies, Inc. *Vicon Motion Systems,* Website: Viconpeak.com.

Piccardi, M. (2004). *Background subtraction techniques: a review*, in Proceedings of the IEEE Interantional Conference on Systems, Man and Cybernetics, 2004, pp. 3099–3104.

Poppe, E. (2007) *Vision based human motion analysis: An overview.* Computer Vision and Image Understanding. 108, pp. 4–18. Website: Elsevier.sciencedirect.com.

Remondino F. *(2006) Imaged-Based Modeling for object and human reconstruction Ph.D. Thesis, ETH No 16562,* Mitteilungen No 91, ISBN 3-906467-61-9, ETH, Zürich, 2006, pp. 69, 70

Schwameder, H., Muller, E., Lindenhofer, E., DeMonte, G., Potthast,W., Brûggemann, P. (2005) *Kinematic characteristics of the early flight phase in ski-jumping*. In: E. Müller, D. Bacharach, R. Klika, S. Lindinger, and. Schwameder (Eds.), *Science and skiing III*, pp. 381–391. Oxford: Meyer & Meyer Sport.

Schwameder, H., (2008) *Biomechanics research in ski jumping*, pp. 1991–2006. Sports Biomechanics January 2008, 7(1), pp. 114–136.

Shigang Liu, Jiancheng Sun and Jianwu Dang 2008. *A Linear Intersection-Resection Bundle Adjustment Method*. Inform. Technol. J., 7, pp. 220–223.

Sæther, B., Rueslåtten, H., Henningsen, T., 1991: *«Advanced Techniques for Lineament Analysis»*. Presented at the Eigth Thematic Conference on Geologic Remote Sensing, Denver, Colorado, USA, April 29 – May 2, 1991.

Virmavirta, M., Isolehto, J., Komi, P., Bruggemann, G-P., Muller, E. and Schwameder, H., (2005) *Characteristics of the early flight phase in the Olympic ski jumping competition*. Journal of Biomechanics 38 (2005), pp. 2157–2163.

Wang, W., Hu, & Tan,T (2003). *Recent development in human motion analysis*. Pattern recognition no 36, pp. 585–601.

Appendix A Collinearity equations

For the sake of simplicity we assume an ideal camera (asprs 2004, p. 220–222) defined by a principal distance, a principal point to which image coordinates refer, and a camera system rotated and translated in the object system as in Fig. A1. This means that the use of this camera model in practice (e.g. for bundle triangulation) assumes linear and non-linear image errors disregarded or eliminated. (The general equations, see (Luhmann et al. 2006, Eq. (8.4), p. 205))

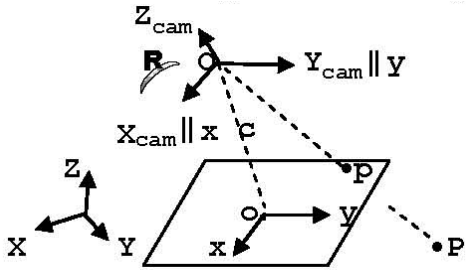


Figure A1 Geometry of the ideal camera model. Projection centre (O). Image point (p). Object point (P). Principal distance (c). Image system (x,y). Principal point (o). Object system (X,Y,Z). Camera system $(X,Y,Z)_{cam}$*. Rotation (**R**) of the camera system.*

The collinear equations are then

$$x = -c\,(\mathbf{X} - \mathbf{X}_0)^T\,\mathbf{r}_1 / (\mathbf{X} - \mathbf{X}_0)^T\,\mathbf{r}_3$$
$$y = -c\,(\mathbf{X} - \mathbf{X}_0)^T\,\mathbf{r}_2 / (\mathbf{X} - \mathbf{X}_0)^T\,\mathbf{r}_3 \qquad \text{(A1)}$$

with:

$\mathbf{X} = (X,Y,Z)^T$, $\mathbf{X}_0 = (X,Y,Z)_0^T$

$\mathbf{R} = (\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) = ((r_{11}, r_{12}, r_{13})^T, (r_{21}, r_{22}, r_{23})^T, (r_{31}, r_{32}, r_{33})^T)$

In homogenous representation:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -c & 0 & 0 & 0 \\ 0 & -c & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \mathbf{R} & -\mathbf{R}\mathbf{X}_0 \\ 0\;0\;0 & 1 \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \end{pmatrix} \qquad \text{(A2a)}$$

or compact

$$\boldsymbol{x} = \boldsymbol{P}\,\boldsymbol{X} \qquad \text{(A2b)}$$

with

$\boldsymbol{x} = (x_1, x_2, x_3)^T$

$\boldsymbol{X} = (X_1, X_2, X_3, X_4)^T$

$\boldsymbol{P}$: 3x4 projection matrix

Notes:
- $\boldsymbol{K}$ = diag(–c,–c,1) is a 3x3 *calibration matrix*
- For a straight-line-preserving camera, $\boldsymbol{K}$ contains the five interior parameters of principal point, principal distance, affinity and shear. With the 6 pose parameters the rank ($\boldsymbol{P}$) = 11. (asprs 2004, p. 224, Eq. 3.139)

– Relations between homogeneous and Euclidian coordinates are:

$$(x_1, x_2)/x_3 = (x, y)$$
$$(X_1, X_2, X_3)/X_4 = (X, Y, Z) \qquad \text{(A2c)}$$

Appendix B Deriving principal distance using vanishing points

The method is taken from Hartley and Zisserman 2003, p. 228.

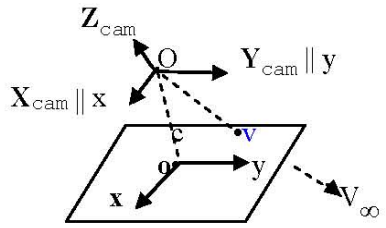


Figure B1 Image system (x, y). Principal point (o). Camera system (X, Y, Z)$_{cam}$. Projection centre (O). Principal distance (c). Vanishing point (v) in image, corresponding point (V) in object

A point vector may be presented in the camera system of the ideal camera (see Figs. B1 and A1) as:

$\mathbf{X}_{cam} = (X, Y, Z)_{cam}^T$: Euclidian coordinate vector

$\boldsymbol{X}_{cam} = (X_1, X_2, X_3, X_4)_{cam}^T = (X, Y, Z, 1)_{cam}^T = (d_1, d_2, d_3, 1)_{cam}^T$: homogenous coordinate vector

where $(d_1, d_2, d_3,)^T$ is the homogenous direction vector of the corresponding projective ray.

If we set $(\mathbf{R}, \mathbf{X}_0) = (\mathbf{I}, \mathbf{0})$ into $\boldsymbol{P}$ in Eq. (A2) and denotes the outcome as $\boldsymbol{P}_{cam}$, we get

$$\boldsymbol{x} = \boldsymbol{P}_{cam}\, \boldsymbol{X}_{cam} \qquad \text{(B2)}$$

Using the notations of Hartley and Zisserman 2003, p. 214, the homogenous vectors of point v and V_∞ are

$$\boldsymbol{v} = (v_1, v_2, v_3)^T \qquad \text{(B3a)}$$

$$\boldsymbol{V}_\infty = (V_1, V_2, V_3, 0)^T = (d_1, d_2, d_3, 0)_{cam}^T \qquad \text{(B3b)}$$

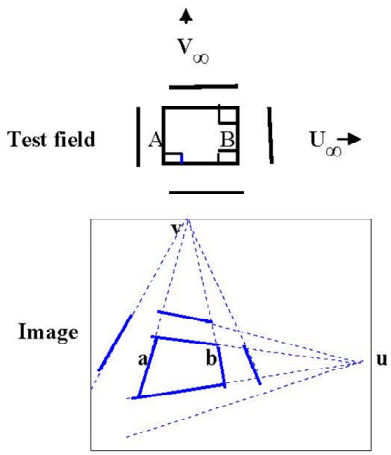


Figure B2 Line featured test field and its image

From (B2) and (B3) we get the fully written matrices:

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} c & 0 & 0 & 0 \\ 0 & -c & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ V_3 \\ 0 \end{bmatrix} = \begin{bmatrix} -c & 0 & 0 \\ 0 & -c & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix}_{V_\infty}$$

or compact

$\boldsymbol{v} = \boldsymbol{K}\, \boldsymbol{d}_{V_\infty}$

or inverted

$\boldsymbol{d}_{V_\infty} = \boldsymbol{K}^{-1}\, \boldsymbol{v}$

Similarly,

$\boldsymbol{d}_{U_\infty} = \boldsymbol{K}^{-1}\, \boldsymbol{u}$

With directions $\boldsymbol{d}_{U_\infty} \perp \boldsymbol{d}_{V_\infty}$ (Fig. B2), i.e.:

$\boldsymbol{d}_{U_\infty}{}^{T}\, \boldsymbol{d}_{V_\infty} = 0$

a constraint on c is obtained

$$\boldsymbol{u}^{T}(\boldsymbol{K}^{-1})^{T}\, \boldsymbol{K}^{-1}\, \boldsymbol{v} = 0 \qquad \text{(B1)}$$

Formula of c is derived from two determined vanishing points u and v (Fig. B2). E.g. vector $\boldsymbol{v}$ is determined from lines $\boldsymbol{a} = (a_1, a_2, a_3)^T$ and $\boldsymbol{b} = (b_1, b_2, b_3)^T$ by cross product $\boldsymbol{v} = \boldsymbol{a} \times \boldsymbol{b}$. Introducing $\boldsymbol{u} = (u_1, u_2, u_3)^T$ and $\boldsymbol{v} = (v_1, v_2, v_3)^T$ into Eq. (B1) the formula of c is derived as

$$c = \sqrt{((u_1\, v_1 + u_2\, v_2)/u_3\, v_3)} \qquad \text{(B2)}$$

Appendix C Point and line intersections in space

The formulation assumes ideal camera as in Fig. A1 and object coordinate system as in Fig. 7. Thus, point intersection is *functionally* modeled by the collinear Eq. (A1) for points; the line intersection by the collinear Eq. (C11) for lines. Eq. (C11) is derived from Eq. (A1) and line equations (C7) and (C8). Input to point and line intersections is i) «error-free» interior/exterior orientations from steps 1 and 2 of flow chart Fig. 4 and ii) observed image coordinates with *stochastic properties* modeled as follows.

With (x, y) representing expectations E(.) of independent coordinate observations (x', y') with errors (ε_x,ε_y), i.e.:

$$x\ (= E(x')) = x' + \varepsilon_x$$

$$y\ (= E(y')) = y' + \varepsilon_y \qquad \text{(C1)}$$

the stochastic properties of (x', y') are modeled as (σ: the standard deviation):

$$\sigma_x = \sigma_y\ (= \sigma_{xy}) \qquad \text{(C2)}$$

Thus, the minimizing constraint for both point and line intersection from an image triplet is simply

$$\Sigma(\varepsilon_x^2+\varepsilon_y^2) \rightarrow \min. \qquad \text{(C3)}$$

Σ over points corresponding to the 3D point (point intersection)

Σ over points on those image lines that correspond to the 3D line (line intersection)

C.1 Point intersection

Assuming Eq. (A1) as the functional model and (C2) as the stochastic model of observed image coordinates x', y' with errors as indicated by (C1), the intersection is formulated as follows:

Eq. (A1), (C1)

↓

$$\begin{aligned}(\mathbf{X}-\mathbf{X}_0)^T\,\mathbf{r}_3\,(x'+\varepsilon_x) &= -c\,(\mathbf{X}-\mathbf{X}_0)^T\,\mathbf{r}_1\\ (\mathbf{X}-\mathbf{X}_0)^T\,\mathbf{r}_3\,(y'+\varepsilon_y) &= -c\,(\mathbf{X}-\mathbf{X}_0)^T\,\mathbf{r}_2\end{aligned} \qquad \text{(C4)}$$

↓

$$\begin{aligned}v_x(=(\mathbf{X}-\mathbf{X}_0)^T\mathbf{r}_3\,\varepsilon_x) &= -c\,(\mathbf{X}-\mathbf{X}_0)^T\mathbf{r}_1-(\mathbf{X}-\mathbf{X}_0)^T\mathbf{r}_3\,x',\ P\\ v_y(=(\mathbf{X}-\mathbf{X}_0)^T\mathbf{r}_3\,\varepsilon_y) &= -c\,(\mathbf{X}-\mathbf{X}_0)^T\mathbf{r}_2-(\mathbf{X}-\mathbf{X}_0)^T\mathbf{r}_3\,y',\ P\end{aligned} \qquad \text{(C5)}$$

where P is the weight given as

$P = ((\mathbf{X}-\mathbf{X}_0)^T\mathbf{r}_3)^{-2}$ with variance factor $\sigma_0^2=\sigma_{xy}^2$, conf. (C2)

With image index 'i' the minimization constraint on «derived» errors (v_x, v_y) is:

$$\Sigma_{i=1,2,3}(P(v_x^2+v_y^2))_i \rightarrow \min \qquad \text{(C6)}$$

The following *algorithm* is based on repeated Least Squares (LSq) estimation

Step a: Set i = 0 and estimate X, Y, Z by LSq as $(X, Y, Z)^i$ from Eq. (C5), constraint (C6) and P = 1

Step b: Set i = i+1 and estimate X, Y, Z by LSq as $(X, Y, Z)^i$ from Eq. (C5), constraint (C6) and P updated with $(X, Y, Z) = (X, Y, Z)^{i-1}$

Step c: If calculated change in estimated X, Y, Z is significant, go to step b

Note: Estimated residuals ε_x, ε_y may be found from Eq. (C4) with estimated X, Y, Z and used to check reliability and derive 3D point accuracy. In the present project 3D point accuracy is rather found by Monte Carlo simulation (Hastedt 2004) of steps 2 and 3 (Fig. 4), assuming standard coordinate deviations in observed image points and in *surveyed reference points,* as reported in 'Observational accuracy', Section 5.

C.2 Line intersection in space

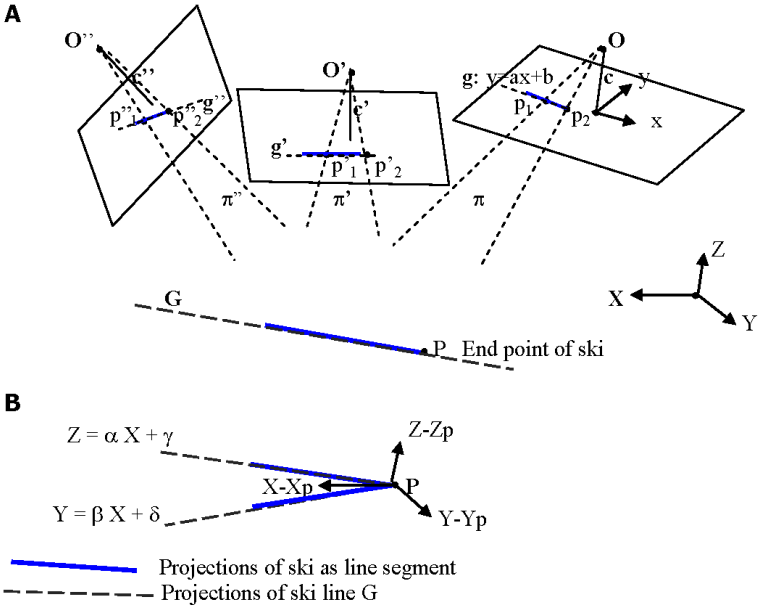


Figure C1 ***A*** *shows a 3D object line G and its projection on three images as 2D lines (g, g' or g"). G is the axis of a pencil of three object planes (π, π', π"). Each plane is defined by the projection center (O, O' or O") and a 2D line. Each 2D line is defined by two arbitrary points (p_i. p'_i or $p"_i$, i = 1, 2).* ***B*** *shows line G projected on X–Z and Y–Z planes.*

With the following equations (C7) and (C8) of image line g and corresponding object line G respectively (see Fig. C1)

$$g: \quad y = a\,x + b \tag{C7}$$

$$G: \quad \begin{aligned} Z &= \alpha\,X + \gamma \\ Y &= \beta\,X + \delta \end{aligned} \tag{C8}$$

and collinearity Eq.(A1) for points , *the collinearity equations of lines* are derived as:

$$a = \mathbf{u}^T\,\mathbf{r}_1 / -\mathbf{u}^T\,\mathbf{r}_2$$

$$b = -c\,\mathbf{u}^T\,\mathbf{r}_3 / -\mathbf{u}^T\,\mathbf{r}_2 \tag{C9}$$

with

$$\mathbf{u} = (\alpha(\delta - Y_0) - \beta(\gamma - Z_0),\ \alpha X_0 + \gamma - Z_0,\ -(\beta X_0 + \delta - Y_0))^T$$

$\mathbf{r}_1$, $\mathbf{r}_2$, $\mathbf{r}_3$ = Column vectors of lines of rotation matrix **R** (see App. A)

In compact form:

$$a = U/W$$

$$b = -c\,V/W \tag{C10}$$

with

$(U, V, W) = (\mathbf{u}^T\mathbf{r}_1, \mathbf{u}^T\mathbf{r}_3, -\mathbf{u}^T\mathbf{r}_2)$

Step A estimates separately for each image line: i) (a, b) as (a', b') and ii) covariance matrix **K** of (a', b') from one error equation (C11) per observed line point. Eq. (C11) is derived as follows

(C1), (C7)

↓

$$v\ (= \varepsilon_y - a\ \varepsilon_x) = -y' + a\ x' + b,\ P \qquad \text{(C11)}$$

where weight P is:

$P = (1+a^2)^{-1}$ with variance factor $\sigma_0{}^2 = \sigma_{xy}{}^2$, conf. (C2)

With index 'j' for n > 2 line points the minimization constraint is

$$\Sigma_{j=1..n}(P\ v_j{}^2) \rightarrow \min \qquad \text{(C12)}$$

Algorithm

1): Set i = 0 and estimate (a, b) by LSq as $(a, b)^i$ from Eq. (C11), P=1 and constraint (C12)

2): Set i = i+1, estimate (a, b) by LSq as $(a, b)^i$ from Eq. (C11), constraint (C12) and P updated with $a = a^{i-1}$

3): By significant change in estimated (a, b) go to (2)

Step B estimates a, b, d, g on the basis of

(i): The two non-linear error equations (C13) per image line

$$a' + \varepsilon_a = U/W$$
$$b' + \varepsilon_b = -c\ V/W \qquad \text{(C13)}$$

(ii): Constraint (C14) on error vector $\varepsilon_{ab} = ((e_a, e_b)_1 .., (e_a, e_b)_m)^T$, m: number of images

$$\varepsilon_{ab}{}^T (\mathrm{diag}(\mathbf{K}_1 .., \mathbf{K}_m))^{-1}\ \varepsilon_{ab} \rightarrow \min \qquad \text{(C14)}$$

Notes

– The diagonal in (C14) is a vector of 2x2 matrices as elements taking into account correlated «observations» a', b'. An alternative is the use of a one-step procedure with one error equation (C15) per line point formulated by setting Eq. (C10) into Eq. (C7), substituting x, y by their expressions (C1) and introducing indices «i» (image line) and «j» (line point):

$$U_i\ \varepsilon x_{ij} - W_i\ \varepsilon y_{ij} = U_i\ x'_{ij} - W_i\ y'_{ij} - c_i\ V_i \qquad \text{(C15)}$$

– For a stereo model singularity occurs if object- and baselines are parallel or object line lies in the epipolar plane.

C.3 Line intersection in homogeneous representation
Notations used (italic font means: «homogenous»):

$\pi = (\pi_1\ \pi_2\ \pi_3\ \pi_4)^T$: Parameters of plane π

$\boldsymbol{g} = (g_1\ g_2\ g_3)^T$: Parameters of image line g

$\boldsymbol{x} = (x_1\ x_2\ x_3)^T = (x, y, 1)^T$: Coordinates of image point p

$\boldsymbol{X} = (X_1\ X_2\ X_3\ X_4)^T = (X, Y, Z, 1)^T$: Coordinates of object point P

$\boldsymbol{N} = (\pi_1\ \pi_2\ \pi_3)^T$: Direction parameters of the normal to plane π

An image point $\boldsymbol{x}^T$ lies on the line g if (Hartley and Zisserman 2003, p. 27)

$\boldsymbol{x}^T \boldsymbol{g} = 0$

Image line g is derived from image points p_1, p_2 as (Hartley and Zisserman 2003, p. 28) :

$\boldsymbol{g} = \boldsymbol{x}_1 \times \boldsymbol{x}_2$

Plane π is derived from line g and projection matrix $\boldsymbol{P}$ as (conf Eq. (A2) and (asprs 2004, p. 237)):

$\pi = \boldsymbol{g}^T \boldsymbol{P} = (\boldsymbol{P}^T \boldsymbol{g})^T$

A point $\boldsymbol{X}$ lies on the plane π if (Hartley and Zisserman 2003, p. 66)

$\pi^T \boldsymbol{X} = 0$

With similar formulae valid for image « **'** » (Fig C1) the condition that a point $\boldsymbol{X}$ lies on the line G is (Hartley/Zisserman 2003, p. 321, section 12.7):

$$\begin{bmatrix} \boldsymbol{P}^T & \boldsymbol{g} \\ \boldsymbol{P}'^T & \boldsymbol{g}' \end{bmatrix} \boldsymbol{X} = \boldsymbol{0}$$

or $\boldsymbol{A}\boldsymbol{X} = \boldsymbol{0}$

Algorithm for finding the «best» image-pair and using this pair for deriving approximate line parameters:
- Form a matrix $\boldsymbol{A}$ with 3 rows of planes π, π',π"and additional, required 4th row of zero-elements
- Do a Singular Value Decomposition of *A* into three 4x4 matrices (Hartley and Zisserman 2003, p. 585, section A4.4):

  $\boldsymbol{A} = \boldsymbol{U}\,\boldsymbol{D}\,\boldsymbol{V}^T$

noticing that columns of $\boldsymbol{U}$ and $\boldsymbol{V}$ are orthogonal, and $\boldsymbol{D}$ is diagonal with singular values.

- Use the two columns of $\boldsymbol{V}$ corresponding to the two largest singular values as the «best» approximation to matrix $\boldsymbol{A}$
- Calculate approximate direction $\boldsymbol{D}$ir(G) using the «best» intersecting planes, e.g. π', π" (Fig. C1):

$\boldsymbol{D}\text{ir}(G) = (\alpha,\beta,1)^T = \boldsymbol{N} \times \boldsymbol{N}"$

- With evaluated α,β introduced in the two equations (C13), solve the unknowns γ and δ

Appendix D Bundle adjustment (BA)

D.1 One-step BA

The one step BA estimates postulated unknowns **U** (vector (D1)) on the basis of a sample of «observations» **L** (vector (D2)) with stochastic properties of **L** expressed as a covariance matrix (D4) and expectations E(**L**) of **L** expressed as functions (D3) of **U**.

$$\mathbf{U} = (\mathbf{X}(\text{unkn})^T, \mathbf{O}^T)^T \qquad \text{(D1)}$$

with

X(unkn): coordinate unknowns of calibration frame points in lab (Fig. 5) and of dynamic/static points in field (Fig. 7)

O: parameters of interior *o*rientation (different for three cameras) and exterior orientation (pose) of camera stations in lab and field.

$$\mathbf{L} = (\mathbf{x}^T, \mathbf{X}(\text{ref})^T)^T \qquad \text{(D2)}$$

with

x: measured coordinates of image points corresponding to calibration frame points and dynamic/static field points,
X (ref): surveyed coordinates of selected field points (Fig. 7)

Noticing that

$$\mathbf{L} = E(\mathbf{L}) + \mathbf{v}$$

where E( ) is the 'expectation' operator and **v** is the error vector, the functional part (A) and the stochastic part (B) of the mathematical model of BA are:

A The expectations E(.) of observations (D2) expressed in terms of unknowns (D1):

$E(\mathbf{x}) = F(\mathbf{U})$	(collinear equations, conf. App. A)	(D3a)
$E(\mathbf{X}(\text{ref})) = \mathbf{X}(\text{unkn})_{\,ref\ points}$	(additional constraints)	(D3b)

B The stochastic properties of observations (D2), defined as a variance matrix K:

$\mathbf{K} = \text{diag}(\text{var}(\mathbf{x})^T, \text{var}(\mathbf{X}(\text{ref})^T) = \text{diag}\,(\sigma_{xy}^2 \ldots, \sigma_{XYZ}^2 \ldots)$
with coordinate variance $\sigma_x^2 = \sigma_y^2 = \sigma_{xy}^2$ of image points and $\sigma_X^2 = \sigma_Y^2 = \sigma_Z^2 = \sigma_{XYZ}^2$ of reference points

The minimization condition is

$$\mathbf{v}^T \mathbf{P}\, \mathbf{v} \rightarrow \min$$

with *weight* matrix

$$\mathbf{P} = \mathbf{Q}^{-1}$$

and *cofactor* matrix:

$$\mathbf{Q} = \mathbf{K}/\sigma_0^2 = \text{diag}\,(1 \ldots, (\sigma_{XYZ}^2/\sigma_0^2) \ldots)$$

with variance factor $\sigma_0^2 = \sigma^2 xy$ conf. (C2)).

Note that matrix **K** (and factor σ_0^2) may be subject to estimation (see App. D.3).

D.2 Free bundle adjustment as applied in the present project
For each iteration step i, add to equation system D(3) the following equations (m = 1..6 for pose parameter and n = 1, 2 … for camera station):

$(E(O'(pose)_{i-1}) = O(pose))_{m,n}$

O(pose): unknown pose

$O'(pose)_{i-1}$: estimated pose from step i–1 as «observation» with guessed accuracy e.g. $\sigma_{\varphi,\omega,\kappa}$= 0.1 gr, $\sigma_{Xo,Yo,Zo}$= 0.5 dm. ($O'(pose)_0$: derived approximate value before BA)

D.3 Estimating observational coordinate accuracy in step 2 (pose estimation)
We will show how to estimate observational image coordinate accuracy σ_{xy} and observational reference coordinate accuracy, σ_{XYZ} assuming $\sigma_x^2 = \sigma_y^2 = \sigma_{xy}^2$ of measured image points (x,y) and $\sigma_X^2 = \sigma_Y^2 = \sigma_Z^2 = \sigma_{XYZ}^2$ of surveyed reference points (X,Y,Z). The minimization condition of BA may be written as

$$Sum = (Sum_{xy} + Sum_{XYZ}) = \mathbf{e}_{xy}^T \mathbf{P}_{xy} \mathbf{e}_{xy} + \mathbf{e}_{XYZ}^T \mathbf{P}_{XYZ} \mathbf{e}_{XYZ} \rightarrow \min \qquad (D4)$$

Sum_{xy} , Sum_{XYZ} : Square sum of weighted residuals in image points and reference points respectively

$\mathbf{e}_{xy}$: vector of $2n_1$ coordinate residuals in n_1 image points

$\mathbf{e}_{XYZ}$: vector of $3n_2$ coordinate residuals in n_2 reference points

$\mathbf{P}_{xy}$= **I** (unit matrix): $2n_1 \times 2n_1$ weight matrix of image coordinates x, y, with variance factor $\sigma_0^2 = \sigma_{xy}^2$

$\mathbf{P}_{XYZ}$ = diag(P_{XYZ}, . .): $3n_2 \times 3n_2$ weight matrix of reference coordinates X, Y, Z with $P_{XYZ} = \sigma_0^2/\sigma_{XYZ}^2$

The estimates of σ_0^2 and P_{XYZ} must fulfill equations (D5) (where Sum and Sum_{XYZ} are the expressions given by (D4))

$$\sigma_0^2 = Sum/r_1 \qquad (D5a)$$
$$\sigma_0^2 = Sum_{XYZ}/r_2 \qquad (D5b)$$

$r_1 = (2n_1 + 3n_2 - u)$: redundancy in image and reference coordinates
u: number of: (i) pose parameters (3 images) and (ii) 3D coordinates of reference and tie points
$r_2 = (n_2 - 7)$: redundancy in reference coordinates (minimum 7 required)

We evaluate iteratively σ_0^2 and P_{XYZ} (needed to further evaluate σ_{xy} and σ_{XYZ} by (D6)) as follows:

- Calculate σ_0^2 by (D5a) after bundle adjustment with initial P_{XYZ} = guessed $\sigma_{xy}^2/\sigma_{XYZ}^2$
- Introduce this value of σ_0^2 into Eq. (D5b) (where Sum_{XYZ} is the expression given by (D4)) and an improved P_{XYZ} is solved.

If the improvement is significant, repeat the bundle adjustment with this P_{XYZ} to find new, improved σ_0 by Eq. (D5a) and new, improved P_{XYZ} by solving Eq. (D5b), a. s. o., else

$$\sigma_{xy} = \sigma_0 , \qquad \sigma_{XYZ} = \sigma_0 / \sqrt{P_{XYZ}} \qquad (D6)$$