On the Theoretical Accuracy of Multi Image Matching, Restoration and Triangulation

Wolfgang Förstner
Institut für Photogrammetrie, Universität Bonn
Nussallee 15, D-53115 Bonn, e-mail: wf@ipb.uni-bonn.de

Abstract

The paper analyses the theoretical precision of integrated multiple image matching and image reconstruction, and the theoretical accuracy of the triangulation from a sequence of images specializing to tri- and binocular stereo. The estimated geometric parameters from multi image matching, used in aerial triangulation for point transfer, turns out to be statistically uncorrelated from the restored image, and the precision of the shift between two images does not depend on the number of images taking part in the multi image matching. Triangulating from an image sequence reveals the variance of the position of points perpendicular to the trajectory to decrease with the number of images whereas the variance of the distance of the 3D-point to the trajectory decreases with the cube of the number of images, taking the distance between the images as given. The case of three images, representative for three line cameras shows the distance to be independent of the central ray.

1 Motivation and Problem

Since aerial images became available at the beginning of this century Photogrammetry geometrically analyzed image sequences. Sensors in remote sensing provide line image sequences, and in the case of three line scanners multiple image sequences. Todays computer power allows the real time analysis of video image sequences.

There are qualitative differences between these technologies: The size of the images, the density of the images, the type and number of spectral channels, the availability and the cost of image data. This leads to quite different application areas, ranging from topographic mapping, to point determination by aerial triangulation and vision based vehicle control to three dimensional mensuration for inspection.

Common to all applications is the need for calibration, orientation and three dimensional reconstruction. The analysis of the geometric properties of this problem has been intensively studied in Computer Vision, leading to deep insight into the structure of the image pair, the image triplet and streams of images, modeled as projective mappings from three dimensional space, and to efficient techniques for the reconstruction of the Euclidean structure of the three dimensional scene (cf. [Spetsakis and Aloimonos 1990], [Faugeras and Robert 1994], [Luong and Faugeras 1996], [Hartley 1997], [Torr and Zisserman 1998], [Triggs 1998], [Avidan and Shashua 1998], [Csurka *et al.* 1998]). This goes along with the development of matching techniques for image sequences which may be dense or sparse (cf. [Horn and Schunck 1998], [Fleet and Jepson 1990], [Beardsley *et al.* 1996]). Here Photogrammetry and Computer Vision are on close paths using all types of correlation, least squares, feature based or relational matching techniques. Differences lie in the perspective: in Computer Vision emphasis is on finding robust techniques which work under quite general conditions, especially in case no approximate values for the parameters of the interior and the exterior orientation are known, accuracy is of secondary interest; in Photogrammetry emphasis is on exploiting the accuracy potential and the prior knowledge about camera parameters, which leads to solutions adequate for applications with well defined boundary conditions, but transferable to other applications only with difficulties. However, both fields are far from reaching the goal of a generally accepted matching technique.

This paper follows the classical photogrammetric flavor. It is motivated by the developments in automatic aerial triangulation and the derivation of Digital Elevation Models from three line cameras, especially the MOMS02-camera. It investigates the accuracy structure of multiple image matching and the resulting accuracy of three dimensional structure.

The first part analyses the precision structure of multiple image matching for multiple point transfer in aerial triangulation which is a generalization of the cross correlation technique for measuring homologous points in image pairs. The second part analyses the accuracy structure of three dimensional points. It is a generalization of the forward intersection of the three line image case. Both topics are closely related to G. Konecny's work (cf. [Konecny 1978, Konecny 1995]).

2 Multiple Image Matching

Template matching is the technique for finding a given signal, the template, in an observed noisy signal, in our context an image. It has several applications in Digital Photogrammetry, e. g. when detecting and locating fiducial marks or control points whose appearance in the image are given by an intensity mask. The only unknown in template matching is the geometric location of the template within the observed image.

The generalization of template matching may proceed in many ways leading to variations, subsumed under the notion of *area based* matching, in contrast to feature based matching:

- allowing for *more complicated geometric transformations* of the template with respect to the image, e. g. an affine transformation (cf. [Ackermann 1984]).
- allowing for *radiometric differences* between template and image in order to reduce the effect of different illumination and film sensitivity. If these

differences are represented as a linear transfer function for the intensities, it is equivalent to normalized cross correlation, the earliest known and most important technique for template matching used in Photogrammtry (cf. [Kreiling 1976, Helava 1976, Konecny 1978]).

- allowing for *more than one image*. Together with the two previous generalization the two-image case is the classical least squares matching approach (cf. [Förstner 1984])
- including *geometric constraints*, e. g. from the orientation of the images (cf. [Grün and Baltsavias 1988])
- simultaneous matching of *several templates* or image patches (cf. [Rosenholm 1987])
- performing *matching in object space* (cf. [Ebner and Heipke 1988, Wrobel 1988, Fua 1995]).

We here want to generalize area based image matching to more than two images and assume the template to be unknown. We want to treat the most simple model, namely a geometric shift. More complicated models can be treated similarly. There will appear no difference between handling two and handling more than two images.

The goal is to investigate the accuracy of the estimated parameters, namely the relative geometric shifts of all images and the reconstruction of the unknown template.

2.1 An Integrated Model for Image Matching and Reconstruction

We assume K images $g_k(x_i, y_i), k = 1, ..., K, i = 1, ..., I$ having I pixels each to be given. They are assumed to be noisy versions of the unknown template $f(x_i, y_i)$ translated by $(x_k, y_k)^T$. The model reads as:

$$g_k(x_i, y_i) = f(x_i - x_k, y_i - y_k) + n_k(x_i, y_i), \quad i = 1, ..., I, \, k = 1, ..., K. \tag{1}$$

This model obviously *integrates image matching and image restoration*, as already required in [Förstner 1984] and realized later by [Ebner and Heipke 1988] and [Wrobel 1988] within surface reconstruction and by [Schenk and Krupnik 1996] within aerotriangulation.

We assume the function $f(x, y)$ to be represented by a discrete grid together with a linear interpolation function. We regularly need values of f and its first derivatives f_x and f_y at non-integer positions, which we assume to be derivable linearly from neighboring values of f, f_x and f_y.

As not all shifts can be determined simultaneously, we impose the restriction

$$\sum_{k=1}^{K} x_k = 0, \quad \sum_{k=1}^{K} y_k = 0 \tag{2}$$

This implicitly fixes the position of the grid of f to the mean position of the given images.

The model eq. (1) is highly nonlinear due to the irregularity of the function f. We assume approximate values to be known with sufficiently high accuracy. They may be determined in a coarse to fine manner and refer to both, the shifts, as well as the unknown template f.

With the approximate values $(x_k^{(0)}, y_k^{(0)})$ and $f^{(0)}(x_i, y_i)$ the linearized model reads as

$$\Delta g_{ik} = -f_{x_{ik}} \Delta x_k - f_{y_{ik}} \Delta y_k + \Delta f_{ik} + n_{ik} \tag{3}$$

with the difference between observed and predicted intensities

$$\Delta g_{ik} = g_k(x_i, y_i) - f^{(0)}(x_i - x_k^{(0)}, y_i - y_k^{(0)}) \tag{4}$$

the partial derivatives

$$f_{x_{ik}} = \left. \frac{\partial f^{(0)}(x, y)}{\partial x} \right|_{(x,y)=(x_i - x_k^{(0)}, y_i - y_k^{(0)})} \tag{5}$$

$$f_{y_{ik}} = \left. \frac{\partial f^{(0)}(x, y)}{\partial y} \right|_{(x,y)=(x_i - x_k^{(0)}, y_i - y_k^{(0)})} \tag{6}$$

the unknown corrections to the shifts, the unknown corrections of the intensities of the template and the noise

$$\Delta x_k = x_k - x_k^{(0)} \tag{7}$$

$$\Delta y_k = y_k - y_k^{(0)} \tag{8}$$

$$\Delta f_{ik} = f(x_i - x_k^{(0)}, y_i - y_k^{(0)}) - f^{(0)}(x_i - x_k^{(0)}, y_i - y_k^{(0)}) \tag{9}$$

$$n_{ik} = n_k(x_i, y_i) \tag{10}$$

We now determine the coefficient matrices $\boldsymbol{A}$ and $\boldsymbol{H}$ of the observation equations and the constraints in the linearized Gauß-Markoff model with constraints

$$\Delta \boldsymbol{b} = \boldsymbol{A} \Delta \boldsymbol{x} + \boldsymbol{e} \tag{11}$$

$$\boldsymbol{0} = \boldsymbol{H} \Delta \boldsymbol{x} \tag{12}$$

where $\Delta \boldsymbol{b}$ constraints the differences between the observations and the predictions, $\boldsymbol{e}$ the observational errors, and $\Delta \boldsymbol{x}$ contains the corrections to the unknown parameters.

For simplicity we assume the images $1, ..., K$ to be of equal size having I pixels. The first two coefficients in (3) are collected in the $I \times 2$ matrix

$$\boldsymbol{A}_{1k} = (\{-f_{x_{ik}} \;\; -f_{y_{ik}}\}) \quad \text{for all } i \in \text{ image } k \tag{13}$$

The third coefficient is collected in the $I \times I$ unit matrix

$$\boldsymbol{A}_{2k} = \boldsymbol{I}_I \tag{14}$$

This results in the full $KI \times (2K + I)$ coefficient matrix, as we have KI observed intensities, $2K$ unknown geometric transformation parameters, namely shifts, and I unknown intensities of the template:

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{0} & \boldsymbol{0} & \ldots & \boldsymbol{0} & \boldsymbol{I}_I \\ \boldsymbol{0} & \boldsymbol{A}_{12} & \boldsymbol{0} & \ldots & \boldsymbol{0} & \boldsymbol{I}_I \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{A}_{13} & \ldots & \boldsymbol{0} & \boldsymbol{I}_I \\ . & . & . & \ldots & . & . \\ \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{0} & \ldots & \boldsymbol{A}_{1K} & \boldsymbol{I}_I \end{pmatrix} \tag{15}$$

The differences between the observations and the predictions are:

$$\Delta \boldsymbol{b} = (\Delta \boldsymbol{b}_k) = ((g_i)_k) \tag{16}$$

With $\Delta \boldsymbol{p}_k = (\Delta x_k, \Delta y_k)^T$ the linearized constraint eq. (2) can be written with 2×2 unit matrices $\boldsymbol{I}_2$

$$\boldsymbol{0} = (\boldsymbol{I}_2 \ldots \boldsymbol{I}_2) \begin{pmatrix} \Delta \boldsymbol{p}_1 \\ \ldots \\ \Delta \boldsymbol{p}_K \end{pmatrix} \tag{17}$$

2.2 The Normal Equations

We therefore obtain the normal equation system $\boldsymbol{N} \Delta \boldsymbol{x} = \boldsymbol{h}$ for the shifts, the unknown template and the two Lagrangian multipliers for the constraints, collected in $\boldsymbol{\lambda}$

$$\begin{pmatrix} \boldsymbol{A}_{11}^T \boldsymbol{P}_{11} \boldsymbol{A}_{11} & \boldsymbol{0} & \ldots & \boldsymbol{0} & \boldsymbol{A}_{11}^T \boldsymbol{P}_{11} & \boldsymbol{I}_2 \\ \boldsymbol{0} & \boldsymbol{A}_{12}^T \boldsymbol{P}_{22} \boldsymbol{A}_{12} & \ldots & \boldsymbol{0} & \boldsymbol{A}_{12}^T \boldsymbol{P}_{22} & \boldsymbol{I}_2 \\ \ldots & \ldots & \ldots & \ldots\ldots & . & . \\ \boldsymbol{0} & \boldsymbol{0} & \ldots & \boldsymbol{A}_{1K}^T \boldsymbol{P}_{KK} \boldsymbol{A}_{1K} & \boldsymbol{A}_{1K}^T \boldsymbol{P}_{KK} & \boldsymbol{I}_2 \\ \boldsymbol{P}_{11} \boldsymbol{A}_{11} & \boldsymbol{P}_{22} \boldsymbol{A}_{12} & \ldots & \boldsymbol{P}_{KK} \boldsymbol{A}_{1K} & \sum_{k=1}^K \boldsymbol{P}_{kk} & \boldsymbol{0} \\ \boldsymbol{I}_2 & \boldsymbol{I}_2 & \ldots & \boldsymbol{I}_2 & \boldsymbol{0}^T & \boldsymbol{0} \end{pmatrix} \cdot$$

$$\cdot \begin{pmatrix} \Delta \boldsymbol{p}_1 \\ \Delta \boldsymbol{p}_2 \\ \ldots \\ \Delta \boldsymbol{p}_K \\ \Delta \boldsymbol{f} \\ \boldsymbol{\lambda} \end{pmatrix} = \begin{pmatrix} \boldsymbol{A}_{11}^T \boldsymbol{P}_{11} \boldsymbol{h}_1 \\ \boldsymbol{A}_{12}^T \boldsymbol{P}_{22} \boldsymbol{h}_2 \\ \ldots \\ \boldsymbol{A}_{1K}^T \boldsymbol{P}_{KK} \boldsymbol{h}_K \\ \sum_{k=1}^K \boldsymbol{P}_{kk} \boldsymbol{h}_k \\ \boldsymbol{0} \end{pmatrix} \tag{18}$$

A practical procedure would reduce this system by the unknown intensities $\Delta \boldsymbol{f}$ of the template, leading to a system with $2K + 2$ unknown parameters, namely the geometric transformation parameters $\Delta \boldsymbol{p}_k$ and the two Lagrangian multipliers. We do not follow this path.

We, however, want to establish a prediction for the precision. We therefore simplify, and assume all $\boldsymbol{A}_{1k}$ to be equal, thus $\boldsymbol{A}_{1k} = \boldsymbol{A}_1$ and $\boldsymbol{P}_{kk} = \boldsymbol{I}_I$. This is reasonable as the $\boldsymbol{A}_{1k}$ contain the derivatives of the template at the grid positions of the given images. In case their grid would be aligned and they would cover exactly the same part of the template, these derivatives would be identical. Assuming the weights to be equal is a standard assumption. With

$$\boldsymbol{N}_0 = \boldsymbol{A}_1^T \boldsymbol{A}_1 = \begin{pmatrix} \sum_{i=1}^I f_{x_i}^2 & \sum_{i=1}^I f_{x_i} f_{y_i} \\ \sum_{i=1}^I f_{y_i} f_{x_i} & \sum_{i=1}^I f_{y_i}^2 \end{pmatrix} \tag{19}$$

omitting the index k in the derivatives, this leads to the following normal equation matrix:

$$\boldsymbol{N} = \begin{pmatrix} \boldsymbol{N}_0 & \boldsymbol{0} & \dots & \boldsymbol{0} & \boldsymbol{A}_1^T & \boldsymbol{I}_2 \\ \boldsymbol{0} & \boldsymbol{N}_0 & \dots & \boldsymbol{0} & \boldsymbol{A}_1^T & \boldsymbol{I}_2 \\ . & . & \dots & . & . & . \\ \boldsymbol{0} & \boldsymbol{0} & \dots & \boldsymbol{N}_0 & \boldsymbol{A}_1^T & \boldsymbol{I}_2 \\ \boldsymbol{A}_1 & \boldsymbol{A}_1 & \dots & \boldsymbol{A}_1 & K\boldsymbol{I}_I & \boldsymbol{0} \\ \boldsymbol{I}_2 & \boldsymbol{I}_2 & \dots & \boldsymbol{I}_2 & \boldsymbol{0}^T & \boldsymbol{0} \end{pmatrix} \tag{20}$$

Using Kronecker notation and with the K-vector $\boldsymbol{e} = (1, ..., 1)^T$ we obtain

$$\boldsymbol{N} = \begin{pmatrix} \boldsymbol{I}_K \otimes \boldsymbol{N}_0 & \boldsymbol{e} \otimes \boldsymbol{A}_1^T & \boldsymbol{e} \otimes \boldsymbol{I}_2 \\ \boldsymbol{e}^T \otimes \boldsymbol{A}_1 & K \otimes \boldsymbol{I}_I & 0 \otimes \boldsymbol{0} \\ \boldsymbol{e}^T \otimes \boldsymbol{I}_2 & 0 \otimes \boldsymbol{0}^T & 0 \otimes \boldsymbol{0} \end{pmatrix} \tag{21}$$

2.3 The Covariance Matrix of the Unknowns

With the K-vector $\boldsymbol{e}^T = (1, 1, ..., 1)$ inversion yields

$$\boldsymbol{N}^{-1} = \begin{pmatrix} (\boldsymbol{I}_K - 1/K{\cdot}\boldsymbol{e}\boldsymbol{e}^T) \otimes \boldsymbol{N}_0^{-1} & \boldsymbol{e} \otimes \boldsymbol{0}^T & 1/K{\cdot}\boldsymbol{e} \otimes \boldsymbol{I}_2 \\ \boldsymbol{e}^T \otimes \boldsymbol{0} & 1/K \otimes \boldsymbol{I}_I & -1/K{\cdot}1 \otimes \boldsymbol{A}_1 \\ 1/K{\cdot}\boldsymbol{e}^T \otimes \boldsymbol{I}_2 & -1/K{\cdot}1 \otimes \boldsymbol{A}_1^T & 0 \otimes \boldsymbol{0} \end{pmatrix} \tag{22}$$

which can be proved by $\boldsymbol{N}^{-1}\boldsymbol{N} = \boldsymbol{I}$.

The upper left 2×2 block matrix contains the weight coefficients of the unknown parameters. This finally yields the covariance matrix of the unknown parameters $\boldsymbol{p}$ and $\boldsymbol{f}$:

$$\boxed{D\begin{pmatrix} \widehat{\boldsymbol{p}} \\ \widehat{\boldsymbol{f}} \end{pmatrix} = \sigma_0^2 \begin{pmatrix} (\boldsymbol{I}_K - 1/K{\cdot}\boldsymbol{e}\boldsymbol{e}^T) \otimes \boldsymbol{N}_0^{-1} & \boldsymbol{e} \otimes \boldsymbol{0}^T \\ \boldsymbol{e}^T \otimes \boldsymbol{0} & 1/K \otimes \boldsymbol{I}_I \end{pmatrix}} \tag{23}$$

Eq. (23) gives insight into the structure of the precision of multi image matching:

- The estimated geometric transformation parameters $\hat{\boldsymbol{p}}$ and estimated signal $\hat{\boldsymbol{f}}$ are uncorrelated. Thus evaluation of the restoration and evaluation of the geometric transformation of the matching process can be performed independently. This results holds for *any type of geometric transformation*, as no use is made of the special structure of the $\boldsymbol{A}_{1k}$ in the derivation.
- Observe the covariance matrix $\boldsymbol{C}_{\widehat{p}\widehat{p}} = \sigma_0^2(\boldsymbol{I}_K - 1/K{\cdot}\boldsymbol{e}\boldsymbol{e}^T) \otimes \boldsymbol{N}_0^{-1}$ of the unknown geometric transformation parameters depends on the inverse normal equation matrix $\boldsymbol{N}_0^{-1}$ known from classical image matching with two images (cf. eq. (19)).

  It seemingly depends on the number K of images. However, the uncertainty of a geometric transformation between any two images k' and k'' in our case of a pure shift is given by

$$D(\boldsymbol{p}_{k'} - \boldsymbol{p}_{k''}) = 2\,\sigma_0^2 \begin{pmatrix} \sum_{i=1}^{I} f_{x_i}^2 & \sum_{i=1}^{I} f_{x_i} f_{y_i} \\ \sum_{i=1}^{I} f_{y_i} f_{x_i} & \sum_{i=1}^{I} f_{y_i}^2 \end{pmatrix}^{-1} \tag{24}$$

due to the correlation between the geometric transformation parameters, resulting from $\boldsymbol{N}_0$ from eq. (19) and

$$D\begin{pmatrix} \boldsymbol{p}_{k'} \\ \boldsymbol{p}_{k''} \end{pmatrix} = \sigma_0^2 \begin{pmatrix} 1 - 1/K & -1/K \\ -1/K & 1 - 1/K \end{pmatrix} \otimes \boldsymbol{N}_0^{-1} \tag{25}$$

Thus more images do not change the precision of the geometric transformation parameters, which appears plausible, as the information of an additional image is used for its own geometric transformation and the improvement of the estimated template. The result coincides with the result from classical least squares matching with two images.

- The precision $D(\hat{\boldsymbol{f}}) = \sigma_0^2/K \cdot \boldsymbol{I}_I$ of the restored image $\widehat{\boldsymbol{f}}$ increases with the number of used images, as to be expected.

The independence of geometric transformation parameters and signal is the most important result of this step: From now on we can assume the signal to be known without loss of generality. The weight matrix $\boldsymbol{P}_{\widehat{p}\widehat{p}}$ of the unknown geometric transformation parameters is $Diag(\boldsymbol{N}_0)$ with the constraint of the geometric transformation parameters to sum to $\mathbf{0}$. Thus we finally have:

$$\boxed{\boldsymbol{C}_{\widehat{p}\widehat{p}} = \sigma_0^2 \left(\boldsymbol{I} - 1/K\, \boldsymbol{e}\boldsymbol{e}^T\right) \otimes \boldsymbol{N}_0^{-1}, \qquad \boldsymbol{P}_{\widehat{p}\widehat{p}} \doteq \boldsymbol{C}^+_{\widehat{p}\widehat{p}} = \sigma_0^{-2}\, \boldsymbol{I} \otimes \boldsymbol{N}_0} \tag{26}$$

independent on the type of geometric transformation and

$$\boxed{\boldsymbol{C}_{\widehat{f}\widehat{f}} = \sigma_0^2/K\, \boldsymbol{I} \qquad \boldsymbol{P}_{\widehat{f}\widehat{f}} = K\sigma_0^{-2}\, \boldsymbol{I}} \tag{27}$$

3 Triangulation in an Image Sequence

Let us now assume we have observed homologous points in an image sequence. We want to investigate the quality of the triangulation of a 3D-point in dependency on the geometric setup. Especially we are interested in the quality of the height or depth for the case of a three line camera, such as the MOMS02 camera. We assume a straight trajectory with the projection centers being of equal distance and the cameras looking perpendicular to the trajectory. This is also the standard case in aerial applications of Photogrammetry.

3.1 The Model

For simplicity we assume the tracked point $P(x, z)$ lies in the xz-plane. The y-coordinate is not taken into account. The point is observed from K cameras with projection centers $O_k(x_{0k}, 0)$ and known fixed orientations. The image points are $P'_k(x_k)$ with standard deviation σ_{x_k}. The principle distance is c. The task is to determine the optimal coordinates $(\widehat{x}, \widehat{z})$.

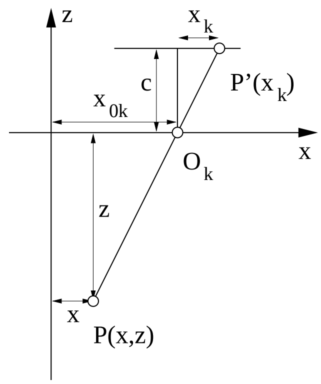
In case the rotation angles are zero the nonlinear model reads as

$$E(x_k) = c \cdot \frac{x - x_{0k}}{z} \qquad D(x_k) = Diag(\sigma_{x_k}^2) = \sigma_0^2\, Diag(p_k^{-1}) \tag{28}$$

With the partial derivatives

$$\frac{\partial x_k}{\partial x} = \frac{c}{z} \quad \text{and} \quad \frac{\partial x_k}{\partial z} = -\frac{c}{z^2}\,(x - x_{0k}), \tag{29}$$

Figure 1: *shows the geometric situation for the k-th camera with projection center O_k.*



where $x = x^{(0)}$ and $z = z^{(0)}$ are the approximate values we obtain the coefficient matrix

$$\boldsymbol{A} = (\boldsymbol{a}_k^T) \quad \text{with} \quad \boldsymbol{a}_k^T = \frac{c}{z^2} \left(z \quad - (x - x_{0k})\right) \tag{30}$$

The normal equation system $\boldsymbol{N}\widehat{\Delta \boldsymbol{x}} = \boldsymbol{h}$ is determined by

$$\boldsymbol{N} = \boldsymbol{A}^T \boldsymbol{P} \boldsymbol{A} = \frac{c^2}{z^4} \begin{pmatrix} z^2 \sum_k p_k & -z \sum_k p_k (x - x_{0k}) \\ -z \sum_k p_k (x - x_{0k}) & \sum_k p_k (x - x_{0k})^2 \end{pmatrix} \tag{31}$$

$$\boldsymbol{h} = \boldsymbol{A}^T \boldsymbol{P} \Delta \boldsymbol{b} = \frac{c}{z^2} \begin{pmatrix} z \sum_k p_k \Delta x k \\ - \sum_k p_k (x - x_{0k}) \cdot \Delta x k \end{pmatrix} \tag{32}$$

with $\Delta x_k = c \cdot \frac{x^{(0)} - x_{0k}}{z^{(0)}} - x_k$. The estimated unknown parameters then are $\widehat{\boldsymbol{x}} = \widehat{\boldsymbol{x}}^{(0)} + \widehat{\Delta \boldsymbol{x}}$.

3.2 The Precision

The *precision* of the unknown coordinates can easily be determined in case we choose the coordinate system such that $\sum_k p_k (x - x_{0k}) = 0$. After inversion of $\boldsymbol{N}$ we obtain the general expression

$$\sigma_{\widehat{x}} = \frac{z}{c} \cdot \frac{\sigma_0}{\sqrt{\sum_k p_k}} \qquad \sigma_{\widehat{z}} = \frac{z^2}{c} \frac{\sigma_0}{\sqrt{\sum_k p_k (x - x_{0k})^2}} \tag{33}$$

In case the images of the sequence have the same distance B, are symmetric with respect to the unknown point and have equal weight, thus $x_{0k} = k \cdot B$, $k = -\frac{K-1}{2}, \ldots, \frac{K-1}{2}$, $x = x^{(0)} = 0$ and, $p_k = p = \sigma_0^2 / \sigma_x^2$, we get

$$\sigma_{\widehat{x}} = \frac{z}{c} \cdot \frac{\sigma_x}{\sqrt{K}} \qquad \sigma_{\widehat{z}} = \frac{z^2}{c \cdot B} \cdot \frac{\sigma_x \cdot \sqrt{12}}{\sqrt{K(K^2 - 1)}} \tag{34}$$

The precision $\sigma_{\widehat{x}}$ of $\widehat{x}$ thus increases with $\sqrt{K}$ while the precision $\sigma_{\widehat{z}}$ increases with $\sqrt{K^3}$.

E. g we obtain for $K=2$, $K=3$ and $K=5$ the following expressions for $\sigma_{\widehat{z}}$:

$$\sigma_{\widehat{z}}^{[2]} = \frac{z^2}{c\,\cdot\,B}\cdot\sqrt{2}\sigma_x \qquad \sigma_{\widehat{z}}^{[3]} = \frac{z^2}{c\,\cdot\,B}\cdot\frac{\sigma_x}{\sqrt{2}} \qquad \sigma_{\widehat{z}}^{[5]} = \frac{z^2}{c\,\cdot\,B}\cdot\frac{\sigma_x}{\sqrt{10}} \tag{35}$$

the first expression being the classical of binocular stereo. The second expression gives the theoretical precision for the three line camera, assuming all three images having the same resolution.

We now want investigate the case $K=3$ of the MOMS02-camera and specialize the equations such that the central image has a different resolution than the other two. We therefore assume $p_1 = p_3 = p_e$, and $p_2 = p_k$, the indices e and i standing for *e*xterior and *i*nterior points or rays, and $x_{01} = -x_{03} = -B$, $x_{02} = 0$ and obtain

$$\widehat{x} = x^{(0)} + \frac{z}{c}\frac{p_e(\Delta x_1 + \Delta x_3) + p_k \Delta x_2}{2p_e + p_k} \qquad \widehat{z} = z^{(0)} - \frac{z^2}{cB}\frac{\Delta x_1 + \Delta x_3}{2p_e} \tag{36}$$

and thus

$$\sigma_{\widehat{x}} = \frac{z}{c}\cdot\frac{\sigma_0}{\sqrt{2p_e + p_k}} \qquad \sigma_{\widehat{z}} = \frac{z^2}{c\cdot B}\cdot\frac{\sigma_0}{\sqrt{2p_e}} \tag{37}$$

Obviously $\widehat{z}$ is completely independent of $x_2 = x_k$, thus on the interior ray!

In the case of the MOMS02-camera we have the following values: $z = 296$ km, $B = 116$ km, $c = 0.237$ m, for the oblique (exterior) stereo channels 6 and 7. We assume a conservative value for the matching accuracy $\sigma_0 = 1/3$ [pel]. This leads to a standard deviation of:

$$\sigma_{\widehat{z}} = \frac{(296[km])^2}{0.237[m]\cdot 116[km]}\cdot\frac{3.3[\mu m]}{\sqrt{2{\cdot}1}} = 7.4[m] \tag{38}$$

This is the Cramer-Rao bound for the precision, stating that the actually reachable precision cannot be better. The theoretical standard deviations 9 to 12 [m] reported in [Schneider and Hahn 1995] are derived by rigorous error propagation include all orientation errors and are only slightly larger than the crude approximation with the above mentioned simplified model. The theoretical values for the precision are confirmed by the empirical standard deviations of appr. 10 [m].

3.3 The Reliability

In order to evaluate the quality of the estimated coordinates we also investigate the reliability, i. e. the checkability of the observations and the sensitivity of the result with respect to gross matching errors [Baarda 1967, Baarda 1968, Förstner 1987].

The *checkability* of the observations needs the redundancy numbers [Förstner 1979]

$$r_k = (1 - \boldsymbol{a}_k^T \boldsymbol{N}^{-1} \boldsymbol{a}_k)\cdot p_k = 1 - \frac{p_k}{\sum_k p_k} - \frac{p_k x_{0k}^2}{\sum_j p_j x_{0j}^2} \tag{39}$$

which in case of equal weights lead to

$$r_k = 1 - \frac{1}{K} - \frac{x_{0k}^2}{\sum_k x_{0k}^2} = 1 - \frac{1}{K} - \frac{12}{K(K^2-1)}\left(k - \frac{K-1}{2}\right)^2 \tag{40}$$

of the k-th ray in a sequence of K images.

The lower bound for outliers detectable with a statistical test then is given by

$$\nabla_0 x_k = \sigma_{x_k} \delta_0 \sqrt{\frac{1}{r_k}} \tag{41}$$

where δ_0 depends on the significance number α of the test and the required power β_0 of the test. We choose $\delta_0 = 4$ in the following.

For the cases $K = 2$, $K = 3$, and $K = 5$ we obtain the redundancy numbers $r_k^{[K]}$:

$$\begin{aligned} r_1^{[2]} &= r_2^{[2]} = 0 & (42) \\ r_1^{[3]} &= r_3^{[3]} = 1/6, \qquad r_2^{[2]} = 2/3 & (43) \\ r_1^{[5]} &= r_5^{[5]} = 2/5, \qquad r_2^{[5]} = r_4^{[5]} = 7/10 \qquad r_3^{[5]} = 4/5 & (44) \end{aligned}$$

For binocular stereo ($K = 2$) we obviously obtain $r_k = 0$, thus the observations, i. e. the x-coordinates are not checkable. In case of three rays all observations are checkable, however no localization is possible. The residuals in the exterior rays 1 and 3 are *always* four times smaller than the residual in the interior ray. The checkability in the case $K = 5$ is much better. Up to two errors in the x-coordinate can be located.

If we, in the case $K = 3$, again distinguish interior and exterior points and assume $p_1 = p_3 = p_e$ and $p_2 = p_i$, we obtain the redundancy numbers

$$r_e = \frac{1}{2} - \frac{p_e}{2p_e + p_i} \quad ; \quad r_i = 1 - \frac{p_i}{2p_e + p_i} \tag{45}$$

In case of the MOMS02-camera we may assume the matching accuracy to be approximately proportional to the pixel size thus $p_i = 9\,p_e$, as the pixel are 4.5 m and 13.5 m for the interior and the exterior ray resp. This leads to the redundancy numbers

$$r_e = \frac{1}{2} - \frac{1}{2+9} = \frac{9}{22} = 0.41 \quad ; \quad r_i = 1 - \frac{9}{2+9} = \frac{2}{11} = 0.18 \tag{46}$$

Observe the checkability of the exterior rays increased due to the higher precision of the central ray.

The boundary values for detectable errors, assuming $\sigma_{x_k} = 1/3$ [pel] are

$$\nabla_0 x_e = 1/3[pel]\, 4\sqrt{\frac{22}{9}} = 2.08[pel] \quad ; \quad \nabla_0 x_i = 1/3[pel]\, 4\sqrt{\frac{11}{2}} = 3.12[pel] \tag{47}$$

Thus matching errors in flight direction need to be 2 or 3 pixels in order to be detectable with a statistical test. Due to the large pixels of the exterior rays the interior ray is less checkable .

The *sensitivity* of the result with respect to outliers is given by the maximum influence of non detectable outliers onto the result. Here we are only interested in the effect of matching errors onto the height. When partitioning the design matrix $\boldsymbol{A}$ into $\boldsymbol{A} = (\boldsymbol{B} \quad \boldsymbol{C})$, where $\boldsymbol{B} = (\boldsymbol{b}_k)$ collects the coefficients for the parameters not taken into account we generally have

$$\nabla_{0k}^{[K]} z = \sigma_z \delta_0 \sqrt{\frac{1 - r_k - u_{t_i}}{r_k}} \tag{48}$$

with

$$u_{t_i} = \boldsymbol{b}_k^T (\boldsymbol{B}^T \boldsymbol{P} \boldsymbol{B})^{-1} \boldsymbol{b}_k \; p_k \tag{49}$$

This in the general case of K views, thus $\boldsymbol{B} = c/k \cdot \boldsymbol{e}$, with *equally* weighted observations, therefore $u_{t_i} = 1/K$, leads to

$$\nabla_{0k}^{[K]} z = \sigma_z \delta_0 \sqrt{\frac{1 - r_k - 1/K}{r_k}} \tag{50}$$

for the k-th ray in a sequence of K images. This yields

$$\begin{aligned}
\nabla_{01}^{[2]} z &= \infty && (51)\\
\nabla_{01}^{[3]} z &= \nabla_{03}^{[3]} z = 6.9\, \sigma_z^{[3]}, \qquad \nabla_{02}^{[3]} z = 0 && (52)\\
\nabla_{01}^{[5]} z &= \nabla_{05}^{[5]} z = 4.0\, \sigma_z^{[5]}, \qquad \nabla_{02}^{[5]} z = \nabla_{04}^{[5]} z = 1.15\, \sigma_z^{[5]}, \qquad \nabla_{03}^{[5]} z = 0 && (53)
\end{aligned}$$

This again shows the weakness of binocular stereo, the insensitivity of the height with respect to errors in trinocular stereo and the inefficiency of the interior ray for determining the height.

For the special case of the MOMS02-camera we need to take the different weights into account and obtain:

$$\nabla_{01}^{[3]} z = \nabla_{03}^{[3]} z = \sigma_z^{[3]} \delta_0 \sqrt{\frac{2p_e + p_i}{p_i}} = 4.4\, \sigma_z^{[3]}, \qquad \nabla_{02}^{[3]} z = 0 \tag{54}$$

The effect of matching errors in flight direction onto the height of the 3D-point is limited to 2 times its standard deviation. Obviously the more precise interior ray, errors in which have no effect onto the height, decreases the sensitivity of the result by a factor of 1.5.

Obviously a simple *algebraical* and *numerical* analysis of the quality of a geometric setup may give clear insight into the role of the individual observations for determining the unknown parameters.

4 Conclusions

The anlysis of the precision and accuracy of multiple image matching presented here may be used for planning geometric schemes for 3D reconstruction. The stochastic independency of image restauration of the determination of the geometric transformations appears to be the most important theoretical result of this study. The type of quality analysis of the geometry of the multiple forward indetrsection may be easily transfered to similar and possibly more complicated situations.

References

[Ackermann 1984] ACKERMANN, F. (1984): High Precision Digital Image Correlation. In: *39. Photogrammetric Week*. Schriftenreihe des Instituts für Photogrammetrie, Heft 9, Universität Stuttgart, 1984.

[Avidan and Shashua 1998] AVIDAN, S.; SHASHUA, A. (1998): Threading fundamental matrices. In: BURKHARDT, H.; NEUMANN, B. (Eds.), *Computer Vision - ECCV '98*, pages 109–123. Springer, LNCS 1406, 1998.

[Baarda 1967] BAARDA, W. (1967): *Statistical Concepts in Geodesy*, Band 2 der Reihe 4. Netherlands Geodetic Commission, 1967.

[Baarda 1968] BAARDA, W. (1968): *A Testing Procedure for Use in Geodetic Networks*, Band 5 der Reihe 1. Netherlands Geodetic Commission, 1968.

[Beardsley *et al.* 1996] BEARDSLEY, P.; TORR, P.; ZISSERMAN, A. (1996): 3D model acquisition from extended image sequences. In: *Computer Vision, ECCV '96*, pages 683–695. LNCS 1964/1965, Springer, 1996.

[Csurka *et al.* 1998] CSURKA, G.; DEMIRDJIAN, D.; RUF, A.; HORAUD, R. (1998): Closed-Form Solutions for the Euclidean Calibration of a Stereo Rig. In: BURKHARDT, H.; NEUMANN, B. (Eds.), *Computer Vision, ECCV '98*, pages 426–442. Springer, LNCS 1406, 1998.

[Ebner and Heipke 1988] EBNER, H.; HEIPKE, C. (1988): Integration of Digital Image Matching and Object Surface Reconstruction. In: *International Archives for Photogrammetry and Remote Sensing, Vol. 27, Part B*, pages 534–545, 1988.

[Faugeras and Robert 1994] FAUGERAS, O.; ROBERT, L. (1994): What can two Images tell us about a third one. In: *ECCV*, pages 485–492. Springer Verlag, 1994.

[Fleet and Jepson 1990] FLEET, D. J.; JEPSON, A. D. (1990): Computation of Component Image Velocity from Local Phase Information. *International Journal on Computer Vision*, 5(1):77–104, 1990.

[Förstner 1979] FÖRSTNER, W. (1979): Das Programm TRINA zur Ausgleichung und Gütebewertung geodätischer Lagenetze. *Zeitschrift für Vermessungswesen*, 104(2):61–72, 1979.

[Förstner 1984] FÖRSTNER, W. (1984): Quality Assessment of Object Location and Point Transfer Using Digital Image Correlation Techniques. In: *Int. Archives of Photogrammetry & Remote Sensing, Vol. 25, III, Rio de Janeiro*, 1984.

[Förstner 1987] FÖRSTNER, W. (1987): Reliability Analysis of Parameter Estimation in Linear Models with Applications to Mensuration Problems in Computer Vision. *Computer Vision, Graphics & Image Processing*, 40:273–310, 1987.

[Fua 1995] FUA, P. (1995): Stereo Reconstruction from Multiple Views. In: GRÜN, A.; KÜBLER, O.; AGOURIS, P. (Eds.), *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, pages 23–32. Birkhäuser, Basel, 1995.

[Grün and Baltsavias 1988] Grün, A.; Baltsavias, M. (1988): Geometrically Constrained Multiphoto Matching. *Photogrammetric Engineering and Remote Sensing*, 54(5):633–641, 1988.

[Hartley 1997] Hartley, R. J. (1997): Lines und Points in Three Views and the Trifocal Tensor. *IJCV*, 2:125–140, 1997.

[Helava 1976] Helava, U. V. (1976): Digital Correlation in Photogrammric Instruments. In: *International Archives for Photogrammetry, Vol. 23, Part II*, 1976.

[Horn and Schunck 1998] Horn, B. K. P.; Schunck, B. G. (1998): Determining Optical Flow. *Artificial Intelligence*, 17:185–204, 1998.

[Konecny 1978] Konecny, G. (1978): Digitale Prozessoren für die Entzerrung und Bildkorrelation. In: *Symp. über den Einsatz digitaler Komponenten in der Photogrammetrie*. Institut für Photogrammetrie und Ingenieurvermessung, 1978.

[Konecny 1995] Konecny, G. (1995): Satelliten-Fernerkundung und Kartographie. *GIS – Geoinformationssysteme*, 8(2):3–12, 1995.

[Kreiling 1976] Kreiling, W. (1976): Automatische Auswertung von Stereobildern durch digitale Korrekation. *International Archives for Photogrammetry and Remote Sensing, Helsinki*, 23-3, 1976.

[Luong and Faugeras 1996] Luong, Q. T.; Faugeras, O. D. (1996): The Fundamental Matrix: Theory, Algorithms and Stability Analysis. *IJCV*, 17:43–75, 1996.

[Rosenholm 1987] Rosenholm, D. (1987): Multi-point matching using the least-squares technique for elevation of three-dimensional models. *Photogrammetric Engeneering and Remote Sensing*, 53(6):621–626, 1987.

[Schenk and Krupnik 1996] Schenk, T.; Krupnik, A. (1996): Ein Verfahren zur hierarchischen Mehrfachbildzuordnung im Objektraum. *Zeitschrift für Photogrammetrie und Fernerkundung*, 64(1):2–10, 1996.

[Schneider and Hahn 1995] Schneider, F.; Hahn, M. (1995): Automatic DTM Generation using MOMS-02/D2 Image Data. In: Fritsch, D.; Hobbie, D. (Eds.), *Photogrammetric Week '95*, pages 85–98. Wichmann Verlag, Karlsruhe, 1995.

[Spetsakis and Aloimonos 1990] Spetsakis, M. E.; Aloimonos, J. (1990): Structure from Motion using Line Correspondences. *IJCV*, 4:171–183, 1990.

[Torr and Zisserman 1998] Torr, P.; Zisserman, A. (1998): Robust Computation and Parametrization of Multiple View Relations. In: *in Proc. of ICCV*, 1998.

[Triggs 1998] Triggs, B. (1998): Autocalibration from planar scenes . In: Burkhardt, H.; Neumann, B. (Eds.), *Computer Vision – ECCV '98*, pages 89–105. Springer, LNCS 1406, 1998.

[Wrobel 1988] Wrobel, B. (1988): Facets Stereo Vision (Fast Vision) - A New Approach Computer Stereo Vision and to Digital Photogrammety. In: *Intercommission Conference on Fast Processing of Photogrammetric Data, Zürich*, pages 231–258, 1988.