



České vysoké učení technické v Praze
Fakulta stavební
Katedra mapování a kartografie

Image matching and its applications in photogrammetry

Ph.D. thesis

Markéta Potůčková

Tutor: **Prof. Ing. Bohuslav Veverka, DrSc.**

Aalborg, July 2004

Contents

Introduction	3
1. Overview on image matching	5
1.1 Digital image	5
1.1.1 Radiometric properties	7
1.1.2 Geometric properties	9
1.1.3 Orientation and georeferencing	10
1.1.4 Geometric transformations and resampling.....	12
1.2 Image matching	15
1.2.1 Area based methods	16
1.2.1.1 Correlation	19
1.2.1.2 Least squares matching	21
1.2.1.3 Image distance	26
1.2.1.4 Mutual information	27
1.2.1.5 Tests on area based matching	29
1.2.2 Feature based methods	46
1.2.2.1 Interest points	46
1.2.2.2 Edges and regions	50
1.3 Outliers, their detection and removal	53
1.3.1 Data snooping	54
1.3.2 Robust adjustment	55
2. Automatic exterior orientation of aerial images based on existing data sets	59
2.1 Automatic measurement of ground control points	61
2.2 Overview of methods for automatic exterior orientation of aerial images	64
2.2.1 Methods based on information from a digital vector map	64
2.2.2 Methods based on existing orthoimages and DTM	66
2.3 Orientation of a single image based on an existing orthoimage and DTM	68
2.3.1 Goals of the investigation and the method overview	69
2.3.2 Combination of a topographic database and an orthoimage for extracting control information	71
2.4 Test of the method	73

2.4.1 Data set description	73
2.4.2 Steps in the application of the method and results	75
2.5 Quality control of orthoimages of the next generation	86
2.5.1 Automatic comparison with existing orthoimage	87
2.5.2 Comparison with the topographic map	90
2.6 Evaluation of the method and of the results	93
3. Automatic DTM check and correction based on two overlapping orthoimages	96
3.1 Digital terrain models and their quality	96
3.2 Basic concepts of DTM generation from aerial images	100
3.3 Check of DTM by means of overlapping orthoimages	102
3.3.1 Principles of the method	103
3.3.2 Test of the method	104
3.3.2.1 Data set description	104
3.3.2.2 Calculation strategy	107
3.3.2.3 Results - Imagery 1:25 000	108
3.3.2.4 Results - Imagery 1:15 000	117
3.3.3 Evaluation of the method	121
Conclusion	123
References	126
Acknowledgement	132
Appendix A – Spatial resection with robust adjustment	
Appendix B – Area based matching in the subpixel range	
B.1 Correlation function	B1
B.2 Least squares matching	B4
Appendix C – CD	

Introduction

Aerial images have been one of the main sources for acquisition of geospatial information. The replacement of analogue photographs by digital images created a requirement for automating all processes connected to the determination of objects from imagery. During the last decades research concentrated on developing procedures for automatic orientation of aerial images, derivation of digital elevation models, building or road extraction etc. Several methods for finding conjugate points in overlapping images have been developed. Nevertheless, the possibility for further investigations, improvements and applications still exists.

Orthoimages and digital elevation models are two main products of digital photogrammetry. They are important layers of geographic information systems and are used in many applications including mapping, urban planning, telecommunication, road construction etc. They should be up to date. A fast and economic production is demanded. The last but important step in their production must be quality control in order to guarantee data sets free of systematic errors and outliers. The methods of quality control should also be automatic, complete, and reliable.

The thesis addresses some of the problems connected to the automatic measurement in digital images. It concentrates on a solution of practical tasks reflecting the demands of National Mapping Agencies – automatic orientation of aerial images based on existing data sets, automatic check of orthoimages, and automatic check and correction of digital terrain models.

The thesis is divided into three chapters. All chapters are connected with the method used for solving the practical problems i.e. area based matching. Each chapter consists of an introductory part giving an overview of the state of the art, a more practically oriented part describing solutions to the problems including tests on real data, and a concluding part evaluating the achieved results.

The first chapter has a more theoretical character. It starts with a brief description of a digital image, its acquisition and basic properties. The main part deals with image matching, especially area based methods. Different similarity measures as correlation coefficient, image distance, and mutual information are studied as well as relations between them. Several tests are carried

out in order to find the best methods that could be applied in the applications dealing with image orientation and DTM checking. Attention is also paid to the method of least squares matching. The last part of the first chapter describes the robust adjustment methods that have been applied in photogrammetry for handling outliers.

The second chapter deals with the method of automatic orientation of images based on existing data sets, namely an orthoimage, a digital terrain model and a topographic map. The idea of the method was created at Aalborg University at the end of 1990s. The possibilities of improvements of the method regarding the degree of automation and accuracy are investigated. The method is developed for the purpose of orthoimage production. Therefore checking the orthoimage of the next generation is included.

The quality control of the DTM is the topic of the third chapter. It concentrates on the method of the DTM check and correction based on finding horizontal parallaxes in two overlapping orthoimages. Attention is paid to the development of procedures that guarantee reliability of the applied method. The goal is to divide an area of the evaluated digital terrain model into a part where the method can be applied and the model can be improved and the second part where other check methods must be used.

The problems that are discussed in the thesis cover the area of digital photogrammetry, image processing and statistics. Most of practical calculations are carried out by means of scripts developed for this purpose in MATLAB[®] v. 6.5.

1. Overview on image matching

Measuring conjugate points in two or more images is one of the most common tasks in photogrammetry. It is a part of procedures like relative and absolute orientations of stereopairs, aerotriangulation, or DTM generation. While in analogue and analytic photogrammetric production it is an operator who measures all points manually, in digital systems this task is supposed to be solved automatically. In literature the process of automatic finding corresponding entities in several images is referred as image matching or as the correspondence problem (Schenk, 1999).

The main part of this chapter gives an overview of basic methods for automatic measurement in digital images especially area and feature based matching. The chapter starts with the definition of a digital image, its acquisition and properties. The last part deals with the computation methods that have been developed to handle outliers in the sets of measurement. It is an important issue in the automatic measurement in images, which is not free from erroneous observations, as it will be shown later.

1.1 Digital Image

A **digital image** can be acquired in two basic ways – directly by using a digital camera or a sensor or indirectly by scanning an analogue photograph. Information about objects displayed in the image is then obtained by analysing or further processing of this data. While a sensitivity of an emulsion and a film development have an influence on a quality of a photograph, the quality of a digital image depends on parameters of CCD (charge-coupled device) chips – photosensitive parts of scanners, digital cameras and sensors. Number of elements in the chip, their size, shape, spectral sensitivity and charge transfer efficiency are the main characteristics of CCD sensors. There are two processes involved in the image acquisition that are closely connected to the mentioned parameters – sampling and quantizing. **Sampling** means discretizing in space. The whole image is divided into ‘picture elements’ or **pixels** (pel), which size and shape depend on the size and shape of capacitors in CCDs. Thus, sampling determines geometric properties of an image. **Quantizing** means assigning the intensity value to a pixel and defines radiometric properties of the image (Schenk, 1999).

An easy data transfer or large spectral sensitivity are examples of advantages of digital cameras and sensors. In accordance with the latest development, they are overtaking the place of film

based cameras. Nevertheless, this process is not so quick due to their relatively high cost especially for airborne applications. Regarding geometric resolution of aerial cameras, more information can still be obtained from the analogue photograph (≈ 60 lp/mm) than from the image taken by a digital sensor (≈ 42 lp/mm or $12 \mu\text{m}$ pixel size corresponds to DMC from Z/I Imaging). All imagery used in practical applications described in chapters 2 and 3 are analogue photographs scanned with a photogrammetric scanner. They were taken at the end of 1990s when high resolution airborne sensors like Z/I Imaging's Digital Mapping Camera (DMC) or Airborne Digital Sensor ADS40 from Leica Geosystems were still in a phase of development and testing.

From the point of view of the applications described later, the digital image is presented as a matrix \mathbf{I} consisting of $r = 1, \dots, R$ rows and $j = 1, \dots, C$ columns. Elements of the matrix carry intensity values. Depending on the type of the image, the matrix consists only of one layer (a grey tone image) or several layers (coloured, multispectral, and hyperspectral images), as shown in Fig. 1. A colour table is an alternative form of an image description.

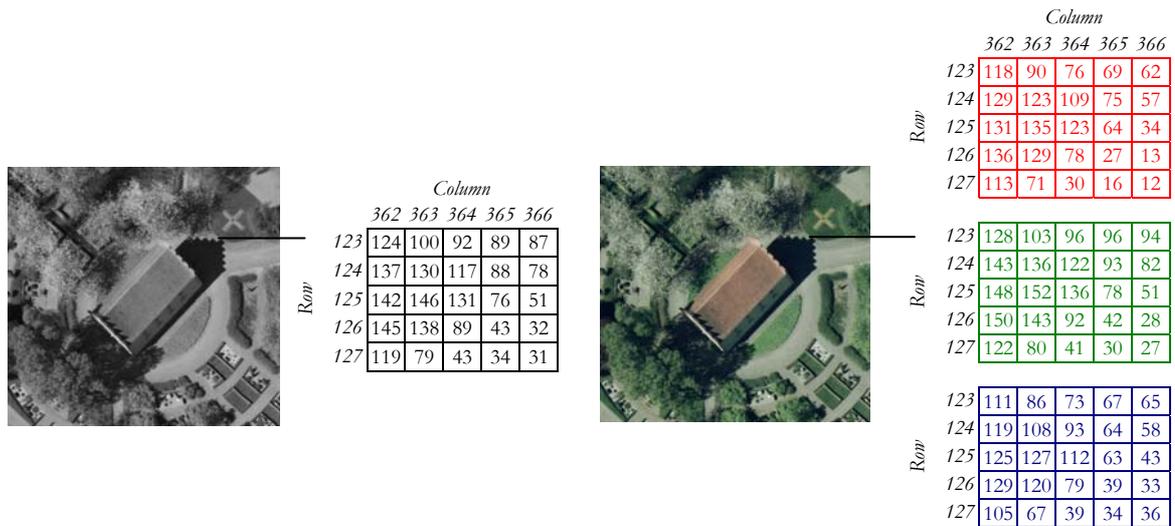


Fig. 1.1 Examples of black & white and coloured images and a matrix presentation of their sections

In the following sections, radiometric and geometric properties of digital images are discussed. An overview of image orientations and geometric transformations follows. The purpose is to set terminology and briefly explain basic terms and processes that are related to the following chapters of this thesis.

1.1.1 Radiometric properties

Number of intensity levels between the minimal (black) and maximal signal (white) and the spacing between levels are parameters of the process of quantizing (Mikhail et al., 2001). The terms **grey values**, image density, or image values are also used in this connection (Jähne, 1997). For practical reasons, 256 levels are the mostly used. An image is then characterised with a radiometric resolution of 8-bit, i.e. an intensity value is represented by an 8-bit number. An optimal number of levels depends on an application. 1-bit images are sufficient for displaying results of operations like edge detection or image segmentation. Higher resolution, e.g. 12-bit or 16-bit images is needed in medicine or remote sensing applications in order to distinguish fine shades of grey. In general, with a higher number of levels the differences to the original signal are smaller as well as an amount of noise introduced into the image.

As mentioned before, only black and white images are used in practical tests of this thesis. An automatic measurement in coloured images can be carried out by converting red, green and blue values into intensity, hue and saturation (Mikhail et al., 2001). All the calculations can then be performed only in intensity e.g. grey tone image. Therefore it is relevant to use the term grey values in the rest of the text although the term intensity values could be used as well.

A **mean grey value** g_m together with a **standard deviation** σ_g (see formula 1.1.) are two statistical characteristics giving information about brightness and contrast in the image (Schenk, 1999).

$$g_m = \frac{1}{RC} \sum_{r=1}^R \sum_{c=1}^C g(r,c) \qquad \sigma_g = \sqrt{\frac{\sum_{r=1}^R \sum_{c=1}^C (g(r,c) - g_m)^2}{RC - 1}} \qquad (1.1)$$

- g_m mean grey value
- σ_g standard deviation (contrast)
- R, C number of rows and columns in the image
- $g(r,c)$ grey value of the pixel at the position r, c

Frequency of each grey value in the image can be expressed by a **histogram**. Frequency is seldom equal for all grey values within the whole range of possible values, as the left part of the Fig. 1.2 shows. An image re-mapping function changes an original histogram into a new one that fulfils requirements for image appearance. Linear, exponential or logarithmic functions can be named as examples. All of them improve contrast only in a certain range of

input grey values. The values out of this range are compressed. Histogram equalisation creates a uniform histogram and in that way equals the contrast over the whole image. Figure 1.2 shows results of a linear histogram stretch and histogram equalisation.

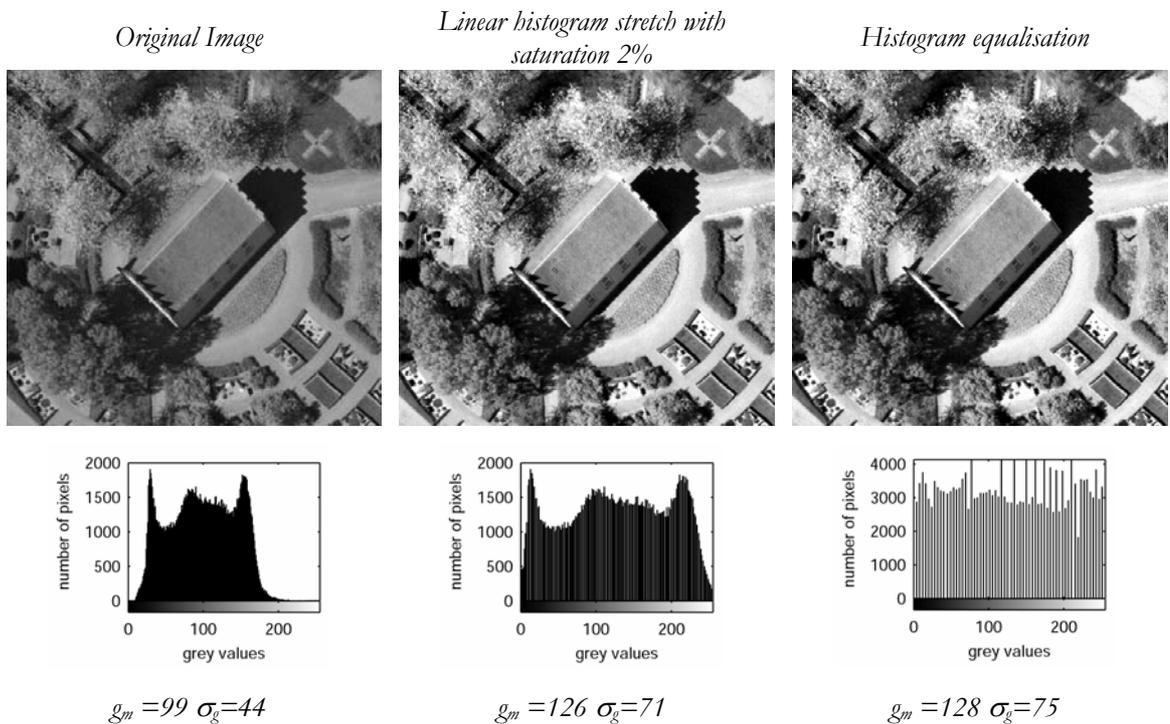


Fig. 1.2 Histogram of an original image and results of a linear histogram stretch and histogram equalisation. Saturation of 2% in this case means that 1% of all pixels with lowest grey values will be assigned value 0 and 1% of all pixels with highest grey values will be assigned maximal grey value 255. g_m and σ_g are the mean grey value and standard deviation (contrast).

Processing of a histogram is useful not only for improving an appearance of an image before a manual measurement or mosaicking of orthoimages but can be also applied before image matching procedures in order to decrease radiometric differences between image patches (see chapter 1.2). Histogram processing of colour images requires a transformation of the red, green, and blue channels into intensity, hue and saturation components (Mikhail et al., 2001).

Entropy H characterises the uncertainty of a grey value in a digital image. It is equal to a number of bits needed for saving a grey value of one pixel. It is calculated by formula 1.2.

$$H = - \sum_{i=0}^{g_{max}} p_i \log_2 p_i \tag{1.2}$$

p_i probability of occurrence of a grey value g_i in the image

g_{max} maximal grey value in the image

Example: It is visible from the histogram of the left image in Fig. 1.2 that the grey values approximately range from 10 to 240. The entropy of this image is $H = 7.3$ which means that in order to prevent all information contained in the image it should be saved with radiometric resolution of 8 bit.

1.1.2 Geometric properties

The size and shape of picture elements are the basic geometric properties of a digital image. The **pixel size** (or geometric resolution) is one of factors having an influence on an accuracy of the measurement. First, in order to be able to recognise an object from a random noise in the image, the object has to cover 2-3 pixels. For getting an overview which level of detail is recognisable in the image, the **ground sample distance** (gsd) is calculated. It is equal to the size of pixel projected on the ground and it can be obtained by multiplication of the pixel size with the scale of the image. Secondly, the experiments showed that accuracy about $1/2 - 1/3$ of a pixel could be achieved with manual measurement (Kraus, 1997). Automatic measurement of signalised ground control points yields accuracy of $1/5$ of a pixel (Hahn, 1997). In close range applications, the accuracy up to $1/1000$ pixel is quoted, depending on a quality of the target and an applied technique (Luhmann, 2000). A **shape of a pixel** is another parameter that plays a role in calculations with a digital image. A square is an often used shape of image grids. A rectangle is characterised by the aspect ratio – a ratio between a pixel width and pixel height. Sensors with hexagonal grids also exist but they have not been implemented in air-born sensors or photogrammetric scanners.

The pixel size together with the radiometric resolution determines the amount of data contained in the image. The geometric resolution of an ordinary aerial photograph corresponds to 72 lp/mm (Kraus, 2000). In order to prevent all information in a corresponding digital image, the scanning must be done with pixel size of $7 \mu\text{m}$. Considering a standard format of aerial photographs $23\text{cm} \times 23\text{cm}$ and a radiometric resolution of 8-bits, the size of an image file will reach 1 GB. Even though a storage capacity and speed of computers are high, such large images can slow all the calculation processes down.

For the purpose of an automatic measurement in images but also e.g. quick zooming, it is advantageous to save an image not only in its original pixel size but in a form of an image pyramid, i.e. together with images of a reduced geometric resolution, so called overviews. The pixel size is reduced by a specified sampling factor f_s . After adding an overview, a size of an original image file S increases according to the geometric series $S(1+1/f_s^2+1/f_s^4+1/f_s^8+\dots)$. It

makes higher requirements to a storage space. When reducing the pixel size, new false details should not be introduced into an overview. Therefore **smoothing** of the image by means of the Gaussian filter is recommended prior to reducing the pixel size (Mikhail et al., 2001).

1.1.3 Orientation and georeferencing

The determination of object co-ordinates of points measured in the images is a basic task of photogrammetry. The geometric relation between an image and object co-ordinate system (**central projection**) is shown in Fig. 1.3. It is analytically described by **collinearity equations** 1.3.

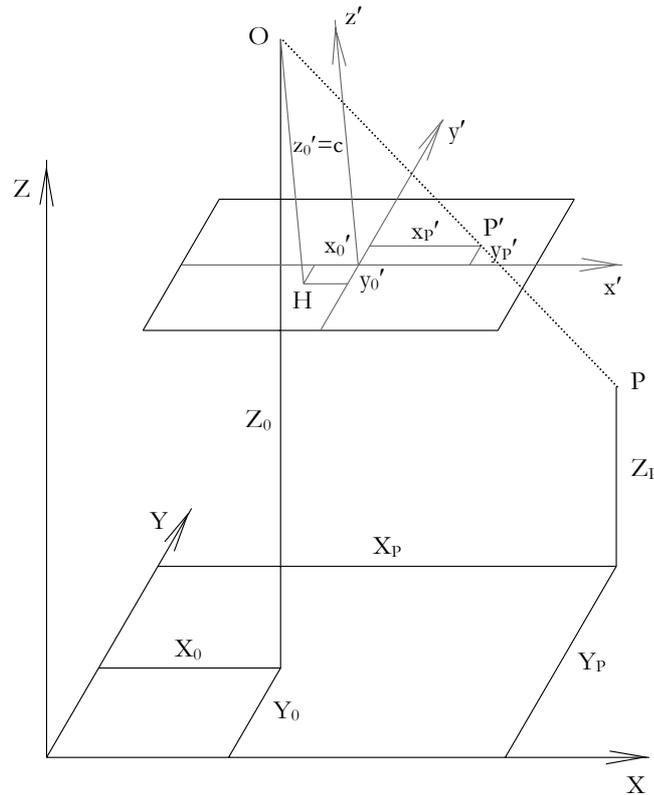


Fig. 1.3 Central projection. Relation between image co-ordinates x_p', y_p' and object co-ordinates X_p, Y_p, Z_p of a point P. The plane given by axes x' and y' of the image co-ordinate system is identical with the image plane. $z'_0=0$ for all points measured in the image. See equation 1.3 for explanation of other symbols.

Collinearity equations:

$$\begin{aligned} x'_p &= x'_0 - c \frac{r_{11}(X_p - X_0) + r_{21}(Y_p - Y_0) + r_{31}(Z_p - Z_0)}{r_{13}(X_p - X_0) + r_{23}(Y_p - Y_0) + r_{33}(Z_p - Z_0)} = x'_0 - c \frac{U}{W} \\ y'_p &= y'_0 - c \frac{r_{12}(X_p - X_0) + r_{22}(Y_p - Y_0) + r_{32}(Z_p - Z_0)}{r_{13}(X_p - X_0) + r_{23}(Y_p - Y_0) + r_{33}(Z_p - Z_0)} = y'_0 - c \frac{V}{W} \end{aligned} \quad (1.3)$$

- x_p', y_p' image co-ordinates of a point P
- X_p, Y_p, Z_p ...object co-ordinates of a point P
- X_o, Y_o, Z_o ...object co-ordinates of the perspective centre O
- r_{ij} elements of the rotation matrix containing angles ω , ϕ and κ (Kraus, 2000)
- c principle distance (camera constant)
- x_o', y_o' image co-ordinates of the principal point H

Parameters of **interior orientation** (image co-ordinates of a principle point, principle distance and lens distortion) and **exterior orientation** (object co-ordinates of a perspective centre and image rotations) must be solved before further use of images for e.g. mapping purposes or orthoimage production. The parameters of interior orientation are usually taken from the camera calibration report. They can also be determined during the process of exterior orientation (self-calibration). In case of a scanned analogue photograph, a relation between image and pixel co-ordinate systems must be found. An **image co-ordinate system** [O'x'y'] is an orthogonal system that usually has its origo in the principal point of best symmetry, units are mm. A **pixel co-ordinate system** [O₁rc] is also an orthogonal right handed system but its origo is in the upper left corner of the image and the units are pixels. The relation between both systems is shown in Fig. 1.4. All measurements are done in the pixel co-ordinate system. In order to apply corrections for lens distortions and having possibility of a direct use of collinearity equations, it is necessary to recalculate positions of the measured points into the image co-ordinate system. **Affine transformation** is usually applied. Fiducial marks are used as identical points. Their photo co-ordinates are known from a camera calibration report. Their image co-ordinates have to be measured in the digital image manually or automatically based on correlation (see chapter 1.2.2).

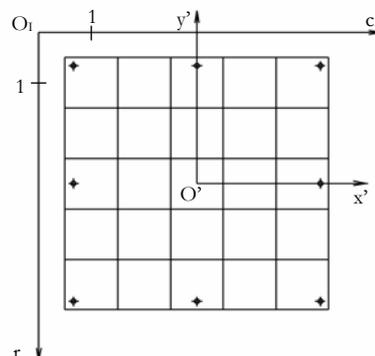
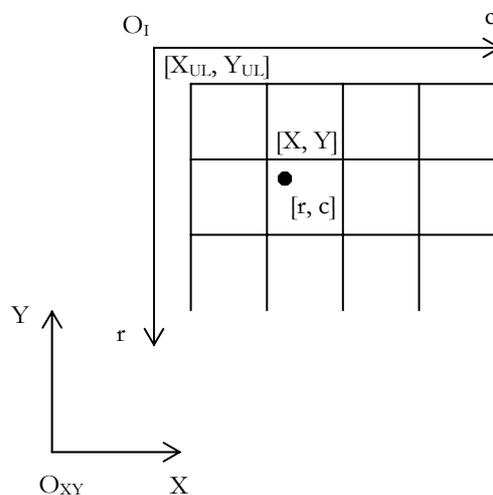


Fig. 1.4 Relation between image and pixel co-ordinate systems. The origo of the pixel co-ordinate system is shifted of a half pixel out from the image in both row and column directions (according to Kraus, 2000).

The techniques for exterior orientation like relative and absolute orientation, aerotriangulation, or spatial resection are well known and developed and can be found in the literature (Kraus, 2000, Mikhail et al., 2001). Possibilities of their automation are mentioned in chapter 2. Spatial resection, that solves exterior orientation of a single image and that is used as an orientation method in the practical application in chapter 2, is described in detail in Appendix A.

A process of finding a position of an image in a reference co-ordinate system is generally called **georeferencing**. In case of photogrammetric products as orthoimages that are often combined with other data sets e.g. topographic maps, the information about their placing in a reference system is often saved in a form of a transformation matrix in an ASCII file connected to an image. In case of a 2D reference system and squared pixels three transformation parameters are only necessary as it is shown in Tab. 1.1.

<i>Parameter</i>	<i>Transformation equation</i>
$X_{UL}=542771.875\ m$	$X = X_{UL}+(c-0.5)gsd$ $Y = Y_{UL}-(r-0.5)gsd$
$Y_{UL}=6250031.250\ m$	$c, r \dots$ co-ordinates of a pixel in the image system
$gsd=0.625\ m$	$X, Y \dots$ co-ordinates of a pixel in the reference (map) system



Tab. 1.1 Transformation between an image and ground reference system. X_{UL} and Y_{UL} are co-ordinates of the upper left corner of the image in the reference system. The last parameter is the ground sample distance (gsd).

1.1.4 Geometric transformations and resampling

An original image is usually taken from a general position but most of applications require an image in a specific position or an image cleared from certain kinds of distortions. Orthoimages

or normalized images (created during automatic exterior orientation or DTM generation) can be used as examples. Geometric transformations of original images are then performed. Due to rotations, scaling and shifts of the original image, the size and shape of pixels of a new image do not correspond to the original values (see Fig. 1.5).

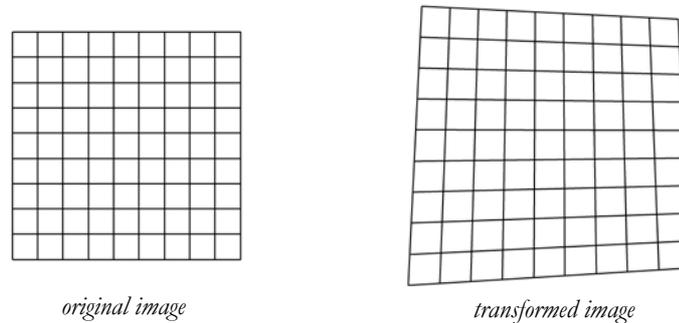


Fig. 1.5 An image grid before and after transformation

Because a regular grid is easier to handle in calculations, a reverse process is applied. An ‘empty’ regular grid overlaying a transformed image is created. By means of a reverse transformation positions of centres of all pixels of a new grid are found in an original image and grey values are interpolated from neighbouring pixels as shown in Fig. 1.6. This process is referred as **resampling**.

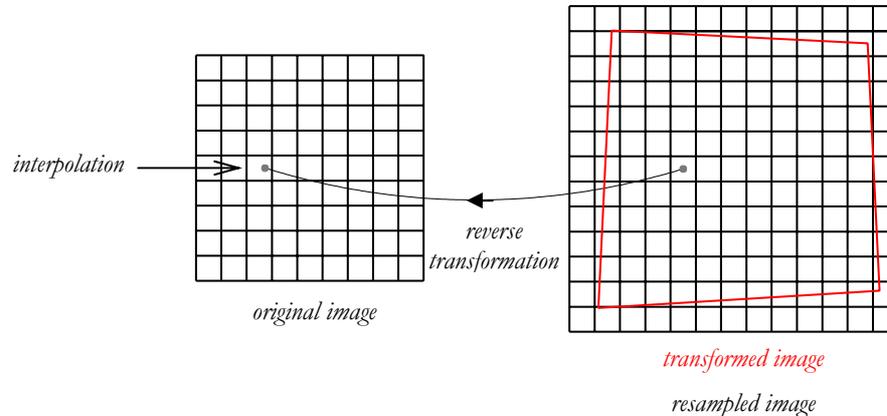


Fig. 1.6 Image resampling. Instead of a transformed image with irregular pixels (a red line shows its border), a new image as an ‘empty’ regular grid is created. The position of each pixel of the new image is found in the original image by means of a reverse transformation. A grey value is interpolated from neighbouring pixels.

The quality of a resampled image depends on an **interpolation algorithm**. Most of the photogrammetric software packages offer following methods:

- **nearest neighbour:** the simplest and fastest method. The distances of a transformed point to centres of neighbouring pixels are calculated. The grey value of the pixel with shortest distance is assigned to the pixel in the new image (Fig. 1.7 left).
- **bilinear interpolation:** Four nearest pixels are taken into account. Distances in row and column directions are used as weights for calculation of a weighted mean (Fig. 1.7 right). The method produces a smoother image than the nearest neighbour method does.
- **cubic convolution:** 16 neighbouring pixels are involved in a calculation. The resulting image is of a better appearance and smoother in comparison to the methods named above.

Practical calculations were performed with orthoimages resampled by bilinear and cubic interpolations and no remarkable differences were observed. The bilinear interpolation was chosen as sufficient for practical calculations presented in the chapters 2 and 3 also due to a shorter calculation time.

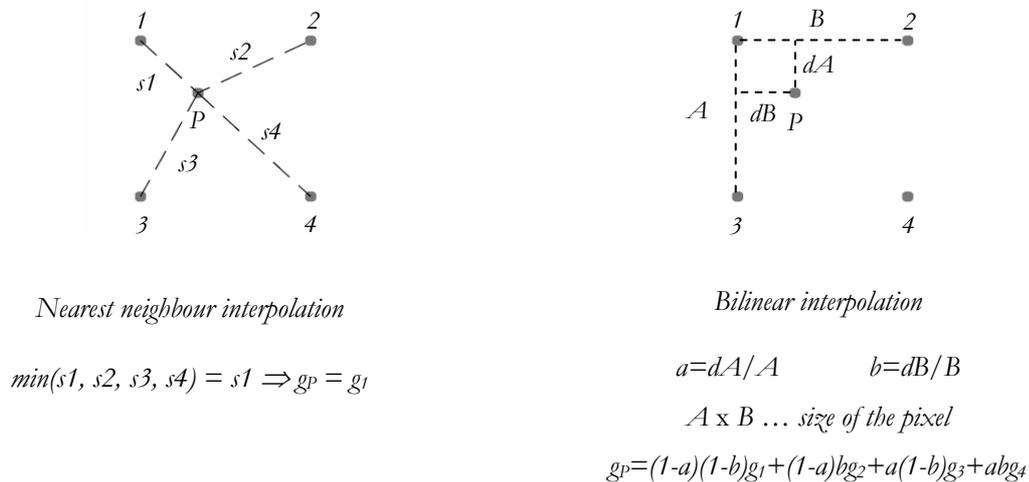


Fig. 1.7 Nearest neighbour and bilinear interpolations. Points 1 to 4 are the centres of pixels in the original image, P is the calculated position of a centre of a pixel of a new image. g_i , $i = 1, \dots, 4$ are known grey values, g_p is an interpolated value.

Tools for orientation and geometric transformation of images are in different levels a standard part of photogrammetric, GIS, or image processing software packages. Some of calculations contained in this thesis were processed by means of

- Z/I Imaging software packages, namely ImageStation™ Digital Mensuration (ISDM), ImageStation™ Base Rectifier (ISBR) and I/RAS C
- own developed or built-in functions of a system MATLAB® and its image processing toolbox

1.2 Image matching

The first solution to the problem of image matching, although analogue in its nature, was given by Hobrough already in late 1950s. A correlator, the first system dealing with automatic finding conjugate points was presented by Wild Heerbrugg company in 1964. The system did not find a wide practical application. Nevertheless, Hobrough's idea of applying cross-correlation found a lot of successors. From early 1970s the development focused on matching digital images and digital correlation was successfully implemented into photogrammetric systems (Helava, 1978). Nowadays, image matching techniques are incorporated into commercial photogrammetric software packages as standard measuring and calculation tools (e.g. ImageStation™ of Z/I Imaging, Match-T and Match-AT of Inpho, Phodis of Zeiss, etc.). In comparison with a manual measurement, automatic methods are faster (especially in aerotriangulation of large image blocks) and the achieved accuracy is in general higher or at least comparable with the accuracy obtained from analytical instruments. On the other hand, due to relatively high amount of mismatches that usually appear a high number of observations (redundancy principle) and implementing techniques for detection and elimination of outliers have an essential importance in order to achieve a high accuracy (Ackerman, 1996a).

Image matching is conventionally performed in the image space. This approach is also an issue of the thesis. The concept of object space matching was also developed (Helava, 1988) but it has not found practical applications so far.

A lot of research has been done with respect to automatic finding tie points connecting two overlapping images (a stereo pair) or connecting images within a block (e.g. Tang and Heipke, 1996, Ackermann, 1996). The search for corresponding points can be done in 2D, e.g. within a rectangle orientated along an approximate epipolar line (see chapter 1.2.1) in case of relative orientation of a stereo pair. In case of known orientation parameters, the search can be restricted only to one dimension i.e. directly on an epipolar line as it is used in the derivation of digital elevation models. Matching between an aerial image and orthoimage of different dates or two orthoimages is also possible and is the issue of the practical tasks described in chapters 2 and 3.

The key issue connected with image matching is a choice of a **matching entity** (a primitive that is compared with a primitive in other images) and a **similarity measure** (a quantitative measure evaluating the match of entities). Matching 'pixel by pixel' over the whole overlapping

area of images means an enormous amount of calculations. Moreover, it leads to ambiguity due to a repetitive occurrence of grey values and due to noise in images. Thus, in general image matching belongs to the group of **ill-posed problems**. It does not fulfil criteria of an existing, unique solution that is stable with respect to small variations in the input data. In order to change image matching into a well-posed problem, such matching entities, similarity measures, geometric constraints, and assumptions must be defined that the space of all the possible solutions will be restricted. Table 1.2 gives an overview of three basic methods of image matching that have been developed and are used in photogrammetry and computer vision.

Matching method	Similarity measure	Matching entity
Area-based	correlation, least-squares matching	grey values
Feature-based	cost function	interest points, edges, regions
Relational	cost function	symbolic description of an image

Tab. 1.2 Image matching methods

In the following sections the matching methods are described in detail. Attention is especially paid to area and feature based techniques. Possibilities of using similarity measures as mutual information and image distance are also discussed.

1.2.1 Area based methods

Grey values are the matching entities in area based matching. Matching one pixel brings an ambiguity problem. Grey values of several neighbouring pixels are therefore used. An image patch cut from one image, so called template, is searched in the second image. The **template** consists of $m \times n$ pixels, mostly $m=n$. The position of the template is referred to the central pixel that is why m and n are odd numbers. The template is compared with patches of the same size in the second image. An approximate position of a corresponding point in the second image can usually be derived (e.g. when approximations of orientation parameters of two overlapping images are known). The comparison is then restricted to an area called a **search area** or window (Schenk, 1999). A value of a similarity measure is calculated at each position of the template within the search area. Depending on the character of the similarity measure, a corresponding point to the centre of the template is assumed to be in the position where maximal or minimal value of the similarity measure is obtained. In photogrammetry, cross-correlation and least squares matching are the mostly used techniques for area based matching. Mutual information and image distance were applied e.g. for registering MR

(magnetic resonance) or CT (computed tomography) images (Maes et al., 1997), in genome engineering (Yu and Jiang, 2001) but also in photogrammetry (Paszotta, 1999). Regardless which of similarity measures is chosen, there are several issues that have to be considered.

Size and location of the template

Larger the template, more the requirement of uniqueness of the matching entity is fulfilled. On the other hand, geometric distortions caused by relief and different orientation of images will influence matching of large templates. The requirement of uniqueness cannot be fulfilled in areas with a repetitive pattern or low contrast and structure. Water or sand areas are typical examples where image matching techniques fail. The areas hidden by high objects should also be excluded. Area based matching as a low level process finds conjugate points in spite of one of them is hidden in one of the images. In similar way, corresponding points on moving vehicles or in shadows lead to incorrect determination of parallaxes. In steep slope areas the corresponding image patches are not geometrically alike. In order to get acceptable results, a size of the template has to be small or its shape adapted to geometric distortion (e.g. a trapezoidal window). One of the possibilities of excluding undesirable areas or finding areas where image matching must be carried out with extra care is using GIS databases. This approach can be easily applied e.g. in DTM generation.

Size and location of the search window

In order to avoid mismatches, the position of the search window must be determined quite accurate in area based matching. Approximations of calculated parameters (e.g. orientation parameters, DTM) and hierarchical approach are usually used for this purpose. Hierarchical approach or coarse-to-fine strategy means that the matching process starts at higher levels of an image pyramid (reduced pixel size) where small details are suppressed. Parameters calculated from the measurements in a higher level of the image pyramid are then used as starting point for matching in a lower level. In the level with the finest geometric resolution the approximations of calculated parameters are good enough for positioning a search window with such accuracy that image matching methods resulting in subpixel accuracy can be applied.

When working with a stereo pair, additional **geometric constraints** can be applied such as epipolar lines. Fig. 1.8 shows a concept of an epipolar line constraint. **Epipolar lines** are intersections of an epipolar plane and image planes. The epipolar plane is defined by projection centres O_1 , O_2 and an object point P . Therefore conjugate points P' and P'' must lie on corresponding epipolar lines e' and e'' . In order to make matching along epipolar lines

easier, images can be transformed to so called normalised images i.e. all epipolar lines in the image are parallel.

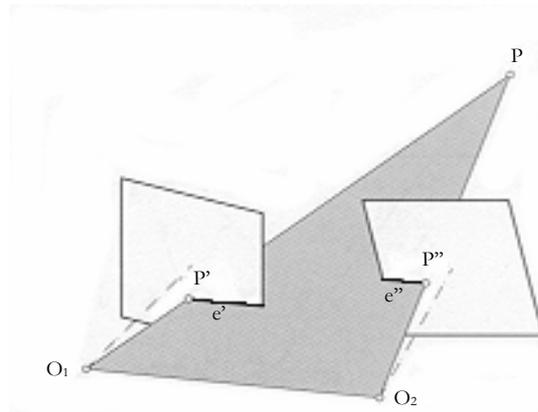


Fig. 1.8 Principle of epipolar geometry. An epipolar plane is defined by projection centres O_1 and O_2 and a object point P . Epipolar lines e' and e'' are intersections of the epipolar plane with the image planes. (adapted from Schenk, 1999)

The size of the search window depends on how precise it is located and on geometric deformations due to relief and image orientations.

Acceptance criteria for the similarity measure

Excluding mismatches is one of the tasks connected to image matching. One possibility how to avoid some of outliers in matching is by setting thresholds for similarity measures. The thresholds can seldom be set for all cases. Although some default values are presented in literature, it happens that the match is successful in spite of exceeding the threshold and vice versa. After the position of 'the best fit' is found, assessment of accuracy and reliability of a found match must be carried out. In addition to thresholding similarity measures, geometrical constrains or robust adjustment techniques are used in further calculations in order to exclude false matches.

1.2.1.1 Correlation

The **normalised cross-correlation coefficient** r is one of common similarity measures used in photogrammetry. It is calculated by the formula 1.4:

$$r = \frac{\sigma_{TS}}{\sigma_T \sigma_S} = \frac{\sum_{i=1}^R \sum_{j=1}^C (g_T(i, j) - \bar{g}_T)(g_S(i, j) - \bar{g}_S)}{\left(\sum_{i=1}^R \sum_{j=1}^C (g_T(i, j) - \bar{g}_T)^2 \sum_{i=1}^R \sum_{j=1}^C (g_S(i, j) - \bar{g}_S)^2 \right)^{1/2}} \quad (1.4)$$

r *normalised cross - correlation coefficient*

σ_T, σ_S *standard deviations of grey values in the template and search image patches*

σ_{TS} *covariance of grey values in the image patches*

g_T, g_S *grey values in the template and search image patches*

\bar{g}_T, \bar{g}_S *mean grey values*

R, C *number of rows and columns of image patches*

If the template and search image patches are represented by vectors $\mathbf{v}_T, \mathbf{v}_S$ of $1 \times RC$ grey values reduced of their means \bar{g}_T and \bar{g}_S , the correlation coefficient can be interpreted as $r = \cos \Theta$, where Θ is an angle between the vectors, as shown in Fig. 1.9 (Kasser and Egels, 2002).

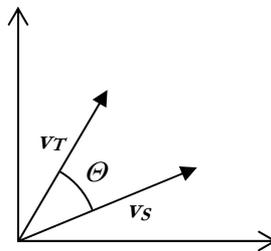


Fig. 1.9 Geometric interpretation of a correlation coefficient $r = \cos \Theta = \mathbf{v}_T \mathbf{v}_S / (|\mathbf{v}_T| \cdot |\mathbf{v}_S|)$

The normalised correlation coefficient has values within the range $-1 \leq r \leq 1$. The value 1 is reached only if image patches \mathbf{g}_T and \mathbf{g}_S are linked by a linear relation $\mathbf{g}_T = r_s \mathbf{g}_S + r_t$, $r_s > 0$, where r_s corresponds to a scale factor and r_t to a shift between grey values in \mathbf{g}_T and \mathbf{g}_S . Values close to 0 indicate no similarity and the value of -1 is obtained when the positive and the negative of an image are matched. Thus, in the image matching process positive values close to 1 are demanded.

A template moves pixel by pixel over the search window and a correlation coefficient is calculated in each position. The position where the correlation coefficient reaches its highest value is selected as a position of the best match/fit (see Fig. 1.10).

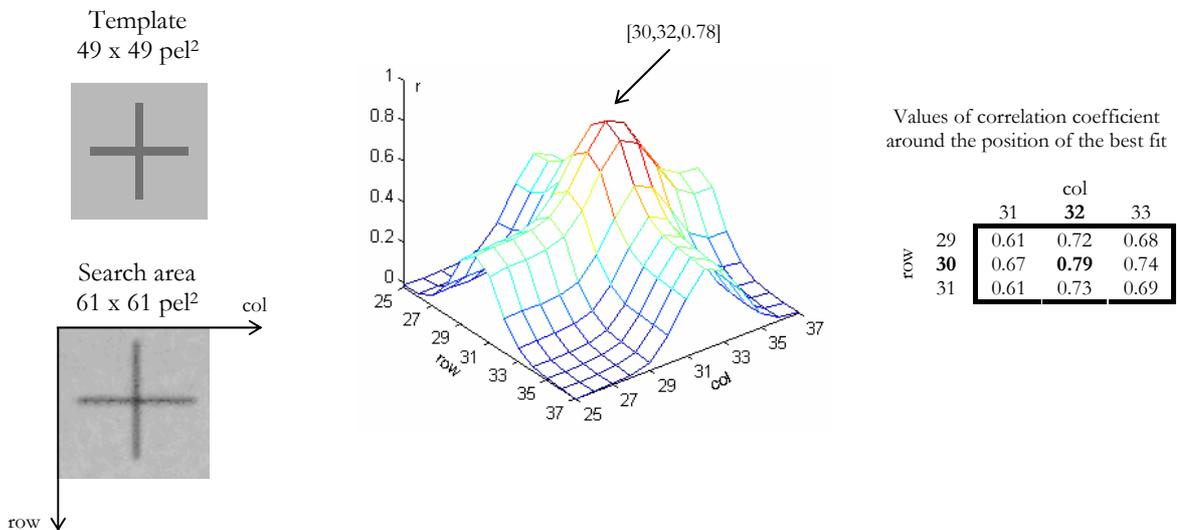


Fig. 1.10 Principle of image matching based on finding maximum of correlation coefficient r . The graph in the middle shows values of the correlation coefficient calculated in 13×13 positions of the template within the search area. The correlation coefficient reaches its maximum 0.79 at the position row=30 and col=32. The search area comes from a photo taken by a réseau camera Rollei 6006 metric that was scanned with pixel size of $21 \mu\text{m}$, the template was created manually.

The correlation coefficient itself does not inform about the accuracy of the found position of the best fit. Some investigations showing a connection between the variance of the determined shift from the centre of the search window, signal to noise ratio, and a size of the template can be found in (Rosenholm, 1986). This theoretical result has not been implemented into practical calculations so far. Nevertheless, it clearly shows that the reliability of a determined position of the best fit depends on radiometric properties of the patches that vary due to different illumination and viewing angle, temporal changes, or projection of matched images. A type of land cover and terrain also plays an important role.

A standard deviation of grey values (formula 1.1) and entropy are measures of contrast and quantity of information in an image patch and they can be used for an evaluation of suitability of the chosen template for matching. **Autocorrelation** can be used for the same purpose as well (see chapter 1.2.1.5, test A). In automated procedures interest or edge operators (see chapter 1.2.2) are applied. In the following matching only those results are accepted where a maximal correlation coefficient exceeds the given threshold. In processes with well

determined objects of measurement like fiducial marks or artificial targets, the thresholding is a successful method for eliminating or at least considerable reducing the number of outliers. E.g. a threshold value of 0.7 proved to be suitable for automatic measurement of fiducial marks. In case of grid crosses or signalized control points when the background is not homogenous, the threshold must be somewhat reduced, e.g. to 0.5 (Kraus, 1997). A similar situation is with natural control points and tie points although in practical applications a value threshold of 0.7 is often a standard. In general, setting a threshold for a correlation coefficient does not mean that all the mismatches are eliminated. When working with natural targets, some good matches have low and some false matches have a high correlation coefficient. By setting a threshold, a number of successful matches are excluded from further calculations while some of mismatches remain. Therefore algorithms for calculating orientation parameters or for DTM generation from matched points must contain routines for eliminating outliers (see chapter 1.3).

If the position of the best fit should be determined with subpixel accuracy, the values of correlation coefficient around its maximum are approximated by a continuous function, e.g. polynomial which parameters are solved in least squares adjustment (Kraus, 1997). The position of the maximum of the polynomial corresponds to the position of the best fit in the subpixel range. Based on standard deviations of polynomial's coefficients derived in least squares adjustment, a standard deviation of the improved position of the best fit can be calculated. In case of searching along an epipolar line a solution is restricted to a correlation curve. The method of approximation of a correlation surface by a 2nd order polynomial including formulas for calculating standard deviations of the derived position is described in detail in Appendix B.1.

1.2.1.2 Least squares matching

Correlation coefficient is not an ideal measure of similarity between two image patches due to their geometric and radiometric differences. At the beginning of eighties development of methods allowing more than cross-correlation like matching in the subpixel range including estimating its accuracy, weighting the observations and blunder detection has started. Among others a method called **least squares matching** (LSM) has been widely investigated. It has found applications both in aerial and terrestrial photogrammetry and has been incorporated in many photogrammetric software packages. Its idea is in minimising differences in grey values between the template and search image patches in an adjustment process where geometric and radiometric corrections of one of matching windows are determined (Schenk, 1999).

The method got an attribute ‘adaptive’ because it gives a possibility of automatic changing the number of parameters and weighting observations depending on their importance and numerical stability of the solution. An important condition for successful LSM is to find an approximate position of the search area relatively accurate, e.g. within a few pixels. Cross-correlation can be used for this purpose.

The relation between grey values of two image patches is expressed by a formula 1.5:

$$g_1(r, c) + e(r, c) = g_2(r, c) \quad (1.5)$$

e is a noise vector caused by different radiometric and geometric effects in both images. In an ideal case of a perfect match $e=0$. The goal is to find such geometric and radiometric transformation parameters of one of the windows, that the vector e is minimised. There are different approaches for choosing a master (i.e. stable, not changing) and a slave (i.e. transformed) window in the matching process. When adapting a slave window, data outside an original patch could be required. A template window is a small, many times artificially made image (e.g. a fiducial mark). Its enlarging would mean extra calculation time. A search window is usually somewhat larger and only its section is used for LSM. Therefore the search window is more suitable for transformation and resampling.

LSM is a non-linear adjustment problem. Due to a **geometric and radiometric transformation** of one of the image patches, the formula 1.5 must be linearized. A solution is then found in an iterative fashion. The linearization of the equation 1.6 formulated for adapting a search window g_S to the template window g_T by means of Taylor’s series is expressed by formula 1.7. This approach published in (Atkinson, 1996, Luhmann, 2000) has been successfully applied in close-range applications. The derivation of formula 1.7 as well as calculating gradients and the adjustment process itself are explained in detail in Appendix B.2.

$$g_T(r, c) + v(r, c) = g_S(f_R(p_1, \dots, p_n, r, c), f_C(p_1, \dots, p_n, r, c))r_s + r_t = \bar{g}_S(r, c) \quad (1.6)$$

$$g_T(r, c) + v(r, c) = g_S^0(r, c) + g_{SR}(r, c)df_R + g_{SC}(r, c)df_C + g_S^0(r, c)dr_s + dr_t \quad (1.7)$$

$$g_S^0(r, c) = g_S(f_R^0(r, c), f_C^0(r, c))r_s^0 + r_t^0$$

$$df_R(r, c) = \frac{\partial f_R(r, c)}{\partial p_1} dp_1 + \dots + \frac{\partial f_R(r, c)}{\partial p_n} dp_n$$

$$df_C(r, c) = \frac{\partial f_C(r, c)}{\partial p_1} dp_1 + \dots + \frac{\partial f_C(r, c)}{\partial p_n} dp_n$$

- r, c row, column
- $g_T(r,c)$ grey values in the template \mathbf{g}_T
- $g_S(r,c)$ grey values in the search area \mathbf{g}_S
- $\mathbf{v}(r,c)$ elements of the vector of residuals \mathbf{v}
- $\bar{g}_S(r,c)$ adjusted grey values in the search area
- $g_S^0(r,c)$ grey values in the search area after applying approximations of geometric and radiometric parameters
- f_R, f_C functions representing geometric transformation between image patches
- p_i geometric parameters
- n number of geometric parameters
- r_s, r_t radiometric scale and shift
- dp_p, dr_p, dr_t corrections to the geometric and radiometric parameters
- g_{SR}, g_{SC} gradients in grey values in row and column directions in the search area

As mentioned in (Kraus, 2000), it may be advantageous to make linearization according to a template window. A design matrix of least squares adjustment then holds stable during all iterations. The linearized equations 1.7 must be modified as follows (equation 1.8):

$$g_S(r,c) + \mathbf{v}(r,c) = g_T^0(r,c) + g_{TR}(r,c)df_R + g_{TC}(r,c)df_C + g_T^0(r,c)dr_s + dr_t \quad (1.8)$$

The number of **geometric parameters** p_i depends on a geometric model used in an adjustment. Image patches cover a relatively small area in the object space. Assuming those areas as planar and a central projection for an image acquisition, a projective transformation fits best. In practical tasks an affine transformation is considered as a sufficient approximation since the image patches to be matched are very small compared to the entire images and formed by a narrow bundle of rays. Tab. 1.3 gives an overview of mostly used geometrical models and a number of unknowns that have to be solved in least squares adjustment.

Due to a different illumination at the moment of taking images, different viewing angles, etc. the radiometric properties of the image patches do not have to be the same. Therefore two **radiometric parameters** namely a shift in brightness r_t and a scale r_s (contrast stretching) are added as unknowns in formulas 1.7 and 1.8. The investigations made by Rosenholm (Rosenholm, 1986) show that the scale r_s better models the radiometric differences between image patches. He concluded that in his tests there was not a significant change in accuracy of LSM if only the scale r_s or both radiometric parameters are included into adjustment. Using

only the shift parameter decreased the accuracy slightly. A problem with using both parameters is in their low convergence ratio due to their high correlation. In order to decrease the number of unknowns and to avoid parameter dependency, radiometric values in both image patches can be adjusted separately prior to LSM (Schenk, 1999). This approach is especially recommended in regions with little texture where the automatic correction of brightness and contrast only amplifies the noise and unsafe little structures (Kraus, 1997). Results of practical tests proving that radiometric adjustment can be carried out prior to LSM are presented in the chapter 1.2.1.5 (test B2).

Least squares matching				First approximation of transformation parameters	
Geometrical model		Transformation equation			
Conform	Translation (2 parameters)	$\begin{pmatrix} r_0 \\ c_0 \end{pmatrix} =$		$t_r^0 = t_c^0 = 0$	
	Translation, Scale (3 parameters)		k	$t_r^0 = t_c^0 = 0$ $k^0 = 1$	
	Translation, Scale, Rotation (4 parameters)		$\begin{pmatrix} a & b \\ -b & a \end{pmatrix}$	$\begin{pmatrix} r \\ c \end{pmatrix} + \begin{pmatrix} t_r \\ t_c \end{pmatrix}$	$a^0 = 1$ $b^0 = t_r^0 = t_c^0 = 0$
	Affine (6 parameters)		$\begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix}$		$a_1^0 = b_2^0 = 1$ $a_2^0 = b_1^0 = t_r^0 = t_c^0 = 0$
Linear radiometric model		$g_S(r,c) = g_S(r,c) r_s + r_t$		$r^0 = 0$ $r_s^0 = 1$	

Tab.1.3 Different geometrical models applied in least squares matching and initial values of unknown transformation parameters.

As mentioned above, the solution of LSM is found in iterations with the first **approximation of transformation parameters** as Tab. 1.3 shows. After each iteration step both the geometric and the radiometric parameters are improved: $\text{par}^{i+1} = \text{par}^i + \text{dpar}^{i+1}$, where par^i is the parameter obtained in i^{th} iteration step and dpar^{i+1} is the correction calculated in $i+1^{\text{st}}$ step. After applying a transformation, the search area changes into an irregular grid and it must be resampled, e.g. by means of bilinear interpolation (see section 1.1.4). Due to the resampling, grey values in the transformed search area become correlated. Nevertheless, this correlation is neglected and the process continues as if the observations were independent. It results in too favourable estimation of variances of unknown parameters, as it will be discussed later. After resampling the computation continues with calculating differences in grey values between the template and the new search area. In case of the formula 1.7 the design matrix of the observation equations is also re-evaluated (see Appendix B.2). Then the next step of the

iteration process starts. Theoretically, the calculation stops as soon as the absolute values of corrections of all transformation parameters are smaller than given limits. The centre of the transformed search image patch after the last iteration is considered as a conjugate point to the centre of the template. In applications carried out in this project, the position of the centre of the template is transformed to the search area and the calculation stops as soon as its change is insignificant (e.g. less than 0.1 pel in both row and column direction for natural targets). A specified maximum number of iterations is also used as a stop criterion for case of slow convergence or divergence.

The geometrical and radiometric models should contain enough parameters to minimise distortions between matched image patches. On the other hand, if some of the parameters cannot be safely determined from the image content the solution becomes numerically unstable or at least the quality of the match deteriorates. Therefore the least squares matching should be adaptive in a sense that its algorithms include testing procedures for excluding non-determinable or unimportant parameters. The tests can be incorporated between individual iteration steps. Because the geometric conditions under which the images were obtained are usually known, a suitable starting set of parameters can be chosen prior to adjustment or estimated from analyses of the image content to be matched (Atkinson, 1996). A size of the matched windows is another parameter that influences an achieved accuracy, convergence, etc. and that can be adapted during the iterative process. If the image patch includes only little detail, it can be enlarged or optionally completely changed automatically in order to cover details that are close but were not included in the previous window. From an accuracy and reliability point of view very small image patches, e.g. $3 \times 3 \text{ pel}^2$ are not suitable for LSM due to small redundancy number. An optimal size of matched windows also differs with a pixel size and scale of the original images. A size of $15 \times 15 \text{ pel}^2$ for good quality images with a lot of details and at least $21 \times 21 \text{ pel}^2$ for noisy images or when significant differences in brightness between a template and a search area occur are recommended in (Kraus, 1997).

Quality of the match is evaluated by means of standard deviations of the derived position of the centre of the template within the search area σ_r and σ_c given by a formula B.11, Appendix B.2. Due to a relatively high number of observations and an iterative process including resampling of a search window, the standard deviations give more optimistic values than a real accuracy is. In practical applications only the **standard deviation of shift parameters** σ_{tr} and σ_{tc} that have the highest influence on values σ_r and σ_c are usually derived for the estimation of

the quality of the match (Atkinson, 1996). Values σ_{tr} and σ_{tc} are calculated by the formula B.12, Appendix B. In literature an accuracy $\sigma_t=(\sigma_{tr}^2+\sigma_{tc}^2)^{1/2}=0.1$ pel to $\sigma_t=0.3$ pel is considered as realistic for non-signalised surface points (Kraus, 1997). An accuracy of σ_t between 0.01 and 0.04 pel is achieved for signalised targets with a good geometry in industrial photogrammetry applications (Luhmann, 2000). The accuracy increases and a convergence gets faster with larger grey-tone gradient. An inflation of the standard deviations of shifts usually appears in cases of slow convergence, i.e. weak or impossible matches.

1.2.1.3 Image distance

Distance represents a simple measure of similarity between two image patches \mathbf{g}_T and \mathbf{g}_S of the same size $R \times C$ pel² and is defined by formula 1.9 (Yu and Jiang, 2001):

$$D(\mathbf{g}_T, \mathbf{g}_S) = \sqrt{\sum_{i=1}^R \sum_{j=1}^C (g_T(i, j) - g_S(i, j))^2} \quad (1.9)$$

From the geometrical point of view the distance D corresponds to the size of a vector $\mathbf{v} = \mathbf{v}_T - \mathbf{v}_S$ where \mathbf{v}_T and \mathbf{v}_S are the vectors representing template and search image patches as shown in Fig. 1.11. The similarity between template and search patches is high if $D(g_T, g_S) \rightarrow 0$.

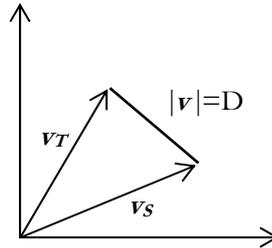


Fig. 1.11 Geometric interpretation of image distance $D = |\mathbf{v}| = |\mathbf{v}_T - \mathbf{v}_S|$

Two modifications of the formula 1.9 can be done in order to reduce an influence of a radiometric shift between patches and of the size of the patches:

- reducing grey values to the mean values \bar{g}_T and \bar{g}_S
- dividing the sum by a number of pixels

The **normalised image distance** D_N is then given by the formula 1.10:

$$D_N(g_T, g_S) = \sqrt{\frac{\sum_{i=1}^R \sum_{j=1}^C ((g_T(i, j) - \bar{g}_T) - (g_S(i, j) - \bar{g}_S))^2}{RC}} \quad (1.10)$$

The relation between the image distance and correlation coefficient (equation 1.11) can be derived directly from the formula 1.10.

$$D_N^2(g_T, g_S) = \frac{\sum_{i=1}^R \sum_{j=1}^C (g_T(i, j) - \bar{g}_T)^2}{RC} - 2 \frac{\sum_{i=1}^R \sum_{j=1}^C (g_T(i, j) - \bar{g}_T)(g_S(i, j) - \bar{g}_S)}{RC} + \frac{\sum_{i=1}^R \sum_{j=1}^C (g_S(i, j) - \bar{g}_S)^2}{RC}$$

$$D_N^2(g_T, g_S) = \sigma_T^2 - 2r\sigma_T\sigma_S + \sigma_S^2 \quad (1.11)$$

Fig. 1.12 shows the relation graphically under an assumption of equal contrast in image patches $\sigma_T = \sigma_S = \sigma$.

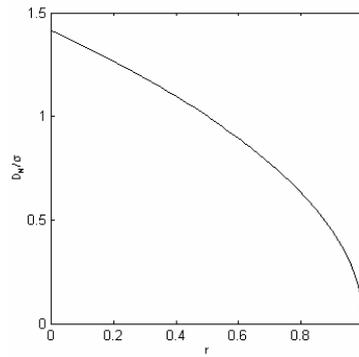


Fig. 1.12 Relation between image distance D_N and correlation coefficient r for $\sigma_T = \sigma_S = \sigma$.

Image matching by means of maximal correlation coefficient and minimal image distance should lead to the same result. In case of well defined targets with no or only minimal distortions in the search area as in the case of the réseau cross in Fig. 1.10 the positions of best fit do not differ. If geometric or radiometric differences exist, the obtained results can differ significantly as will be shown in the chapter 1.2.1.5 (test C).

1.2.1.4 Mutual information

Mutual information is a measure of a statistical dependence between two random variables or a measure of the amount of information that one variable contains about the other (Maes et al., 1997). It is maximal if the image patches are geometrically aligned. The mutual information $I(\mathbf{g}_T, \mathbf{g}_S)$ of two image patches \mathbf{g}_T and \mathbf{g}_S containing grey values in the range of $t=1, \dots, n_T$ and $s=1, \dots, n_S$ is expressed by a formula 1.12.

$$I(\mathbf{g}_T, \mathbf{g}_S) = \sum_{t=1}^{n_T} \sum_{s=1}^{n_S} \hat{p}_{TS}(t, s) \log_2 \frac{\hat{p}_{TS}(t, s)}{\hat{p}_T(t) \hat{p}_S(s)} \quad (1.12)$$

$I(\mathbf{g}_T, \mathbf{g}_S)$ mutual information

\hat{p}_T probability of grey values $t = 1, \dots, n_T$ in the image patch \mathbf{g}_T

\hat{p}_S probability of grey values $s = 1, \dots, n_S$ in the image patch \mathbf{g}_S

\hat{p}_{TS} joint probability

The values \hat{p}_T and \hat{p}_S are obtained by normalisation of image patch histograms. The values of joint probability \hat{p}_{TS} are calculated from a joint image histogram that can be presented as a matrix $n_T \times n_S$ which elements equal to number of pixels where $\mathbf{g}_T(i, j)=t$ and $\mathbf{g}_S(i, j)=s$. An example is given in Tab. 1.4.

		\hat{p}_{TS}								
		\mathbf{g}_T						\mathbf{g}_S		
1	2	3	4	4	1	2	3	3	4	
1	3	4	5	5	2	3	4	4	5	
2	4	6	6	7	2	5	5	6	7	
2	6	6	8	8	1	6	6	7	8	
3	5	7	7	7	3	5	7	7	8	
										\hat{p}_T
t \ s	1	2	3	4	5	6	7	8		
1	0.04	0.04								0.08
2	0.04	0.08								0.12
3			0.12							0.12
4			0.04	0.08	0.04					0.16
5				0.04	0.08					0.12
6					0.04	0.12				0.16
7							0.12	0.04		0.16
8							0.04	0.04		0.08
\hat{p}_S	0.08	0.12	0.16	0.12	0.16	0.12	0.16	0.08		

$$I(\mathbf{g}_T, \mathbf{g}_S) = 1.96$$

Tab. 1.4 Probabilities of grey values \hat{p}_T and \hat{p}_S in image patches \mathbf{g}_T and \mathbf{g}_S , joint probability \hat{p}_{TS} , and mutual information $I(\mathbf{g}_T, \mathbf{g}_S)$.

The relation between the mutual information $I(\mathbf{g}_T, \mathbf{g}_S)$ and the entropy of image patches $H(\mathbf{g}_T)$ and $H(\mathbf{g}_S)$ is calculated according to the formula 1.13 (Maes et al., 1997):

$$I(\mathbf{g}_T, \mathbf{g}_S) = H(\mathbf{g}_T) - H(\mathbf{g}_T | \mathbf{g}_S) = H(\mathbf{g}_S) - H(\mathbf{g}_S | \mathbf{g}_T) = H(\mathbf{g}_T) + H(\mathbf{g}_S) - H(\mathbf{g}_T, \mathbf{g}_S) \quad (1.13)$$

$H(\mathbf{g}_S | \mathbf{g}_T), H(\mathbf{g}_T | \mathbf{g}_S)$... conditional entropy of \mathbf{g}_T given \mathbf{g}_S and of \mathbf{g}_S given \mathbf{g}_T respectively

$H(\mathbf{g}_T, \mathbf{g}_S)$ joint entropy

If image patches \mathbf{g}_T and \mathbf{g}_S are statistically independent, then $\hat{p}_{TS}(t,s)=\hat{p}_T(t)\hat{p}_S(s)$ and $I(\mathbf{g}_T, \mathbf{g}_S)=0$. If \mathbf{g}_T and \mathbf{g}_S are maximally dependent, then $I(\mathbf{g}_T, \mathbf{g}_S) = H(\mathbf{g}_T) = H(\mathbf{g}_S) = H(\mathbf{g}_T, \mathbf{g}_S)$.

In order to find out whether there are differences in found positions of the best fit when using different similarity measures, results from matching based on correlation coefficient and mutual information were compared. Similarly to image distance, there were no differences in case of well-defined targets and a good geometric and radiometric correspondence between templates and search areas (see a cross in Fig. 1.10 as an example). In case of radiometric and geometric differences between image patches the found position of the best fit based on mutual information and correlation coefficient can differ up to several pixels as tests presented in the chapter 1.2.1.5 showed.

1.2.1.5 Tests on area based matching

Several tests on area based matching were carried out. The first test deals with using autocorrelation for checking a suitability of a chosen template for matching. The goal of other tests is to get experience for carrying out calculations connected to the practical applications presented in the chapters 2 and 3 and also to show relations between different similarity measures mentioned in the previous chapters.

The tests are divided into four main categories and several subcategories:

- A) Autocorrelation
- B) Least squares matching
 - B1) Calculation according to formula 1.7 and 1.8
 - B2) Radiometric corrections
- C) Relation between similarity measures
 - C1) Correlation coefficient and image distance
 - C2) Correlation coefficient and mutual information
 - C3) Correlation coefficient, mutual information, and image distance
- D) Combination of similarity measures for the detection of mismatches
 - D1) Position of the best fit calculated for each of similarity measures
 - D2) Position of the best fit calculated by means of cross-correlation; image distance and mutual information are used as attributes

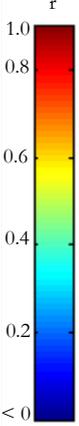
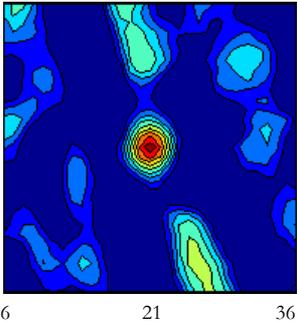
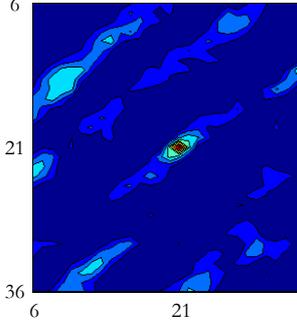
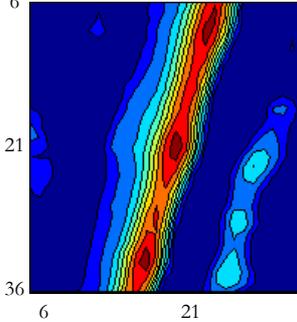
The data used in tests A and B differ and will be described separately. In test C and D, the data described in detail in the chapter 2.4 are used, namely an old and new orthoimages. The words 'old' and 'new' are connected to the fact that there is four years difference in the date of images. Both orthoimages have a pixel size of 0.625 m and were derived from images 1:25 000 taken with a wide-angle camera. The orientation parameters of the old image were derived by

aerotriangulation. The orientation parameters of the new image by means of the method based on an existing orthoimage and a DTM described in the chapter 2. The same DTM was used for derivation of both orthoimages. 122 templates of 31×31 pel² containing road crossings were derived from the old orthoimage and matched in the new one (search area of 61×61 pel²). The described data are chosen because they fit to the application studied in the chapter 2, namely quality control of a new produced orthoimage and partly also to the method of DTM checking that is the topic of the chapter 3.

A) Autocorrelation

Autocorrelation gives a possibility for finding out whether a chosen template is a good candidate for a successful matching. A template and a search area are chosen from the same image. The template is scanned over the search area, a correlation coefficient is calculated at each position and a correlation surface or curve is evaluated. An existence of one sharp peak is demanded. In comparison to the standard deviation of grey values and entropy, not only a template but also its surrounding is involved in the calculation. Therefore it gives a possibility to evaluate a uniqueness of a template within a search area. Examples of three different templates together with the obtained correlation surfaces, values of standard deviation of grey values σ_T (contrast) and entropy H_T are shown in Tab. 1.5.

Both the first and the third templates are characterised with a good structure and relatively high contrast in contrary to the second template. If the second template is matched with a search area that is not cut from an identical but only a similar image (e.g. when matching the left and right image of a stereopair), a few positions with almost identical but very low correlation coefficient would be found probably at the light blue and green areas of the figure showing the distribution of the correlation coefficient values. The third template itself is very suitable for correlation but it is not unique within the search area – three significant maxima appeared in the correlation surface due to the repetitive pattern. Thus, only the first example is desirable for matching.

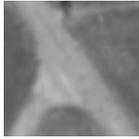
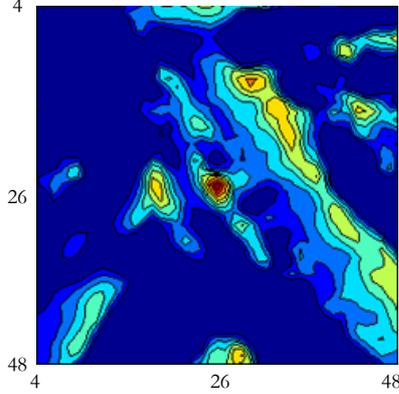
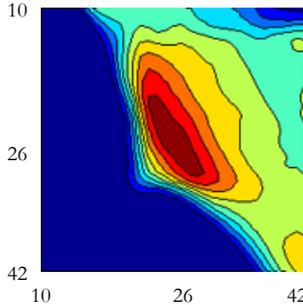
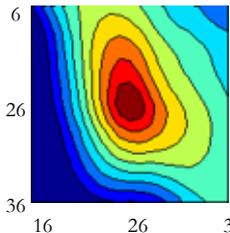
		single object			low contrast			repetitive pattern				
		SA	T		SA	T		SA	T			
												
Correlation surface												
												
			Local maxima				Local maxima				Local maxima	
		row	col	r		row	col	r		row	col	r
	$\sigma_T = 25$				$\sigma_T = 8$				$\sigma_T = 40$	21	21	1
	$H_T = 6$	21	21	1	$H_T = 5$	21	21	1	$H_T = 7$	9	24	0.94
										33	18	0.95

Tab. 1.5 The application of autocorrelation for evaluating suitability of an image patch for matching. The figures in the middle show the distribution of the correlation coefficient. All negative values were changed to 0. The size of the templates (T) is $11 \times 11 \text{ pel}^2$ and the size of the search area (SA) $41 \times 41 \text{ pel}^2$, $1 \text{ pel} \approx 21 \mu\text{m}$. The standard deviation of grey values σ_T and entropy H_T give an overview about contrast and an amount of radiometric information in the template.

It was already mentioned that the size of template and search windows also plays an important role in order to obtain reliable results. Tab. 1.6 shows an example of an application of autocorrelation for finding a proper template size for matching a road crossing.

From the distribution of correlation coefficient values can be concluded that the smallest template $7 \times 7 \text{ pel}^2$ does not fulfil the requirement for its uniqueness within a search window. It does not contain any significant structure, the contrast is very low. It can be well matched in several places along the road. The larger template $19 \times 19 \text{ pel}^2$ brings a much better result but still a lack of structure in its lower left corner causes ellipsoidal shape of the correlation surface around the position of the maximum with the semi-major axis parallel to the road.

The template $31 \times 31 \text{ pel}^2$ seems to have the most suitable size and contrast from the three given examples. Larger templates would not bring remarkable improvements in finding a position of the best fit by means of correlation. Moreover with larger templates the danger of radiometric and geometric differences between images increases. This fact has to be taken into account individually depending on the type of imagery.

SA	 51 x 51pel ²		
T	 7 x 7pel ²	 19 x 19pel ²	 31 x 31pel ²
r	 4 26 48 4 26 48	 10 26 42 10 26 42	 6 26 36 16 26 36
	$\sigma_T = 4$ $H_T = 4$	$\sigma_T = 16$ $H_T = 6$	$\sigma_T = 23$ $H_T = 6$

Tab. 1.6 The use of autocorrelation for finding an appropriate template size. The figures in the middle show the distribution of correlation coefficient. All negative values were changed to 0. $1 \text{ pel} \approx 21 \mu\text{m}$. The standard deviation of grey values σ_T and entropy H_T give an overview about contrast and an amount of radiometric information in the template.

The previous examples showed that autocorrelation can be easily used for checking a suitability of templates and search areas for image matching. The disadvantage is that the obtained results are reliable only on assumption that there are not many differences between search areas derived from tested (i.e. autocorrelated) and matched images. Autocorrelation means that the processing time for image matching doubles. It is probably the reason why it has not been applied in available software packages so far. In automated photogrammetric processes like tie point measurements or DTM derivation, the position of templates is chosen

by means of interest or edge operators (see chapter 1.2.2) that search for points suitable for matching. Nevertheless, incorrect template size or repetitive patterns, e.g. phenomena that can be discovered by autocorrelation, can become a reason for false matches. In spite of its potential, autocorrelation is not used in practical applications described in the chapters 2 and 3 mostly due to long calculation time that is needed for computing area based matching in MATLAB®.

B) Least squares matching

The goal of the test B1 is to find out whether there are any differences when the design matrix of observation equation changes with each iteration (formula 1.7) or remains stable (formula 1.8). The test B2 shows what the difference is when the radiometric adjustment is carried out prior to LSM or when radiometric parameters are included into LSM. All calculations were carried out with own developed MATLAB® functions that can be found in Appendix C.

B1) Calculation according to formula 1.7 and 1.8

In order to find out whether there are any differences in found position of the best fit when using formula 1.7 (transformation parameters found for a search area) or 1.8 (transformation parameters found for a template) two images were created and resampled by means of a conform and affine transformations. The position of the best fit of the centre of the template was calculated by means of known transformation parameters and then derived twice by means of least squares matching. Tab. 1.7 summarises the results.

It can be concluded that there are no significant differences in positions of the best fit obtained by two methods. The calculated standard deviations of shift parameters t_x and t_y do not differ between methods as well. In comparison to reference values, the transformed positions of the centres of the templates do not differ more than 0.04 pel which is acceptable with respect to the quality of search image patches (blurred edges of bright objects due to resampling of an original template from which the search area was created). Method M2 is the one with a stable design matrix. The accuracy achieved by this method is slightly better. On the other hand, more iteration steps are necessary.

Template		Search area	
		Conform transformation $\alpha = 20^\circ$ $k=1.2$	Affine transformation $\alpha = 20^\circ$ $k_r = 0.9$ $\kappa=0^\circ$ $k_c = 1.1$
T1	 25 x 25 pel ²	 45 x 45 pel ²	 42 x 36 pel ²
T2	 35 x 35 pel ²	 59 x 59 pel ²	 51 x 55 pel ²

	Conform transformation		Affine transformation	
Reference	Position of best fit [pel]		Position of best fit [pel]	
	T1: r=23.26 c=22.91 T2: r=29.88 c=29.89		T1: r=18.31 c=21.38 T2: r=25.66 c=27.64	
Calculated	Position of best fit [pel]	Standard deviation of shift parameters [pel]	Position of best fit [pel]	Standard deviation of shift parameters [pel]
T1, M1	r=23.27 c=22.94	$\sigma_{tr}=0.05$ $\sigma_{tc}=0.04$	r=18.31 c=21.40	$\sigma_{tr}=0.06$ $\sigma_{tc}=0.06$
T1, M2	r=23.27 c=22.94	$\sigma_{tr}=0.04$ $\sigma_{tc}=0.04$	r=18.32 c=21.40	$\sigma_{tr}=0.05$ $\sigma_{tc}=0.05$
T2, M1	r=29.91 c=29.92	$\sigma_{tr}=0.10$ $\sigma_{tc}=0.10$	r=25.68 c=27.67	$\sigma_{tr}=0.13$ $\sigma_{tc}=0.12$
T2, M2	r=29.90 c=29.93	$\sigma_{tr}=0.09$ $\sigma_{tc}=0.09$	r=25.67 c=27.66	$\sigma_{tr}=0.10$ $\sigma_{tc}=0.10$

	T1, M1	T1, M2	T2, M1	T2, M2	T1, M1	T1, M2	T2, M1	T2, M2
Number of iterations	14	16	12	25	18	20	14	28

M1: Calculation according to formula 1.7 (transformation parameters for a search window)

M2: Calculation according to formula 1.8 (transformation parameters for a template window)

Tab. 1.7 Comparison of results of least squares matching according to formulas 1.7 and 1.8.

B2) Radiometric corrections

In the second test LSM of 62 road crosses was carried out. The templates were derived from an orthophoto, search areas from an aerial image which was taken five years later (the data set described in the chapter 2.4). Approximate positions of best fit were found by means of the correlation coefficient. Six parameters geometric model of LSM was used. The first calculation included two radiometric parameters into LSM, in the second calculation a radiometric

adjustment was done prior to LSM. Calculations were done according to the formula 1.7. In order to suppress an influence of outliers only matches with number of iterations less or equal to 10 were included into evaluation of results as Tab. 1.8 shows.

	differences '6+0' – '6+2'	
	dr [pel]	dc [pel]
mean [pel]	0.02	-0.06
RMSE [pel]	0.20	0.24
number of points (it ≤ 10)	50	

Mean standard deviations in shift parameters		
	σ_{tr} [pel]	σ_{tc} [pel]
'6+0'	0.27	0.23
'6+2'	0.25	0.22

6+0: LSM with 6 geometric parameters, radiometric adjustment done prior to LSM

6+2: LSM with 6 geometric and 2 radiometric parameters

Tab. 1.8 Comparison of results of least squares matching with two radiometric parameters and radiometric adjustment prior to LSM. Only points with the number of iterations smaller than 11 were taken into the evaluation.

The results show that there is no difference between radiometric adjustment prior to or inside least squares matching. The obtained RMSE values correspond to standard deviations of shift parameters. This conclusion is in an agreement with a research made by Rosenholm (Rosenholm, 1986). The tests B1 and B2 contributed to the following conclusions:

- Calculation according to formula 1.7 (method M1) has only a slightly worse accuracy but a better convergence (compare Tab. 1.7) and therefore all further calculations are based on this formula.
- Due to problems with convergence that can occur because of 'over-parameterisation' (Schenk, 1999), the approach with a radiometric adjustment prior to LSM is chosen for further calculations.

C) Relation between similarity measures

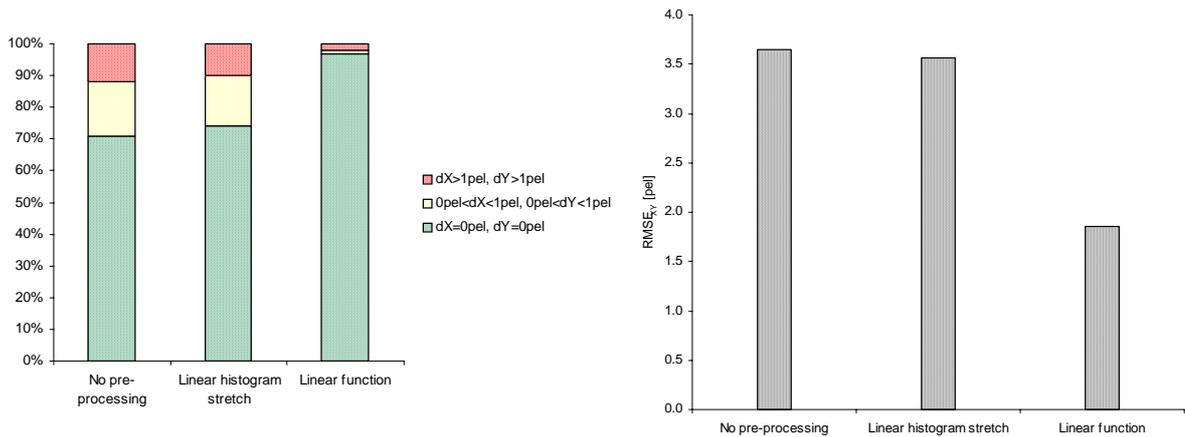
Three similarity measures were described in the previous chapter. The goal of the test C is to find out how the found positions of the best fit obtained by studied similarity measures differ

and why. The image patches of 122 road crosses extracted from two overlapping orthoimages of different date were used for this test as was mentioned at the beginning of this chapter.

C1) Correlation coefficient and image distance

The positions of the best fit between orthoimage patches were derived by means of the normalised correlation coefficient (formula 1.4) and the normalised image distance (formula 1.10). The results were compared. An expectation was that the difference would not exceed 1 pixel (0.625 m) in both X and Y co-ordinates. By subtracting the mean grey value, the radiometric shift between image patches is reduced but no care is taken of differences in contrast. Therefore three calculations with a different level of radiometric adjustment were carried out. The first calculation was done without any image pre-processing. The second one included a histogram stretch of the search area into the range of a template before the distance was calculated. In the third calculation grey values of a search area were transformed by means of two linear radiometric parameters derived in least squares adjustment. Tab. 1.9 summarises the obtained results.

	$dX_{r-D}=0$ pel and $dY_{r-D}=0$ pel	$0 \text{ pel} < dX_{r-D} \leq 1 \text{ pel}$ and $0 \text{ pel} < dY_{r-D} \leq 1 \text{ pel}$	$dX_{r-D} > 1 \text{ pel}$ or $dY_{r-D} > 1 \text{ pel}$	RMSE [pel] dX dY		Mean [pel] dX dY	
No radiometric pre-processing	71% (87 points)	17% (21 points)	12% (14 points)	4.2	3.0	0.5	0.0
Linear histogram stretch	74% (90 points)	16% (20 points)	10% (12 points)	3.8	3.3	0.3	-0.1
Linear function LSA	97% (118 points)	1% (1 points)	2% (3 points)	1.8	1.9	-0.1	-0.1



Tab. 1.9 Differences in the position of the best fit obtained by means of normalised correlation coefficient (r) and normalised image distance (D_N) when different radiometric adjustments of search area were applied prior to the calculation of the image distance. The calculation was done for 122 road crosses derived from two orthoimages. Size of the template was $31 \times 31 \text{ pel}^2$ and size of the search area $61 \times 61 \text{ pel}^2$. $RMSE_{XY} = ((RMSE_X^2 + RMSE_Y^2) / 2)^{1/2}$. $1 \text{ pel} \approx 0.625 \text{ m}$.

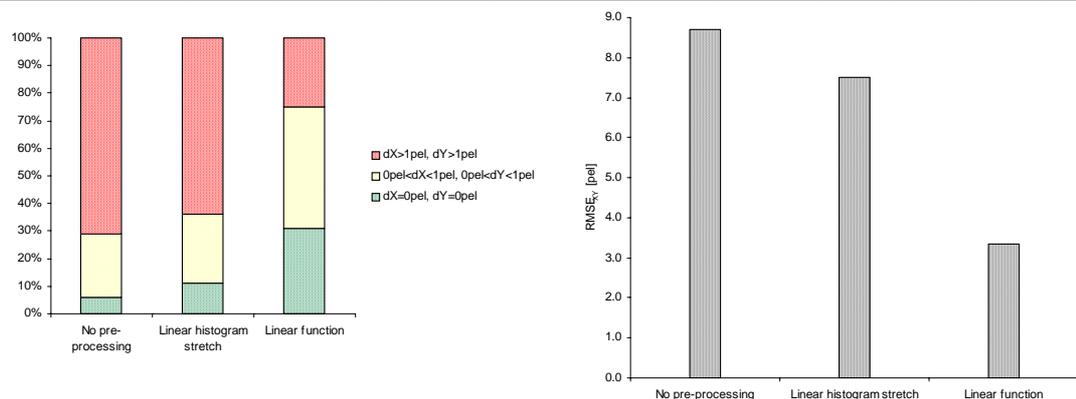
The table shows that when using correlation coefficient and image distance as similarity measures, the obtained positions are identical in most cases if grey values of the matched image patches are adjusted prior to the calculation of the image distance by means of a linear function. Such an adjustment is not necessary in case of the correlation coefficient because the influence of a scale radiometric parameter disappears due to a division of covariance and standard deviations. At three points where the positions differed significantly the conditions for image matching were not ideal; the maximal correlation coefficient did not exceed 0.45.

It can be concluded that image distance is an equivalent measure to the correlation coefficient under the condition of a radiometric adjustment of image patches. Such an adjustment means an extra calculation time. A correlation coefficient approach is therefore simpler and faster and will be used further.

C2) Correlation coefficient and mutual information

Mutual information was also calculated three times applying different radiometric corrections of the search area as in the calculation of the image distance. Differences in found positions of best fit are summarised in Tab. 1.10.

	$dX_{r-1}=0$ pel and $dY_{r-1}=0$ pel	$0 \text{ pel} < dX_{r-1} \leq 1 \text{ pel}$ and $0 \text{ pel} < dY_{r-1} \leq 1 \text{ pel}$	$dX_{r-1} > 1 \text{ pel}$ or $dY_{r-1} > 1 \text{ pel}$	RMSE [pel] $dX_{r-1} \ dY_{r-1}$		Mean [pel] $dX_{r-1} \ dY_{r-1}$	
No radiometric pre-processing	6% (7 points)	23% (28 points)	71% (87 points)	8.4	9.0	0.5	0.8
Linear histogram stretch	11% (13 points)	25% (31 points)	64% (78 points)	7.1	7.9	0.3	-1.2
Linear function LSA	31% (38 points)	44% (54 points)	25% (30 points)	3.2	3.5	0.3	0.0



Tab. 1.10 Differences in the position of the best fit obtained by means of correlation coefficient (r) and mutual information (I) when different radiometric adjustment of search area were applied prior to calculation of mutual information. The calculation was done for 122 road crosses derived from two orthoimages. Size of the template was $31 \times 31 \text{ pel}^2$ and size of the search area $61 \times 61 \text{ pel}^2$. $RMSE_{XY} = ((RMSE_X^2 + RMSE_Y^2) / 2)^{1/2}$.

The differences are somewhat larger than differences between results obtained by means of correlation coefficient and image distance. Radiometric balancing between image patches brought a better agreement in results. However, an expected value of the root mean square error of one pixel in both co-ordinates was exceeded more than three times.

The reason for such differences in obtained positions of the best fit comes from the definition of both similarity measures and the origin of the images and can be explained with help of the examples in Fig. 1.13. The search area was derived from the template in four different ways:

- A unique value was assigned to each original grey value. The relation is linear (in Fig. 1.13 $S=2+3T$).
- A unique value was assigned to each original grey value. The relation is not linear.
- The original grey values were changed only slightly. The assignment is not unique but has a tendency to be linear.
- The original grey values were changed only slightly. The assignment is not unique and does not have a tendency to be linear.

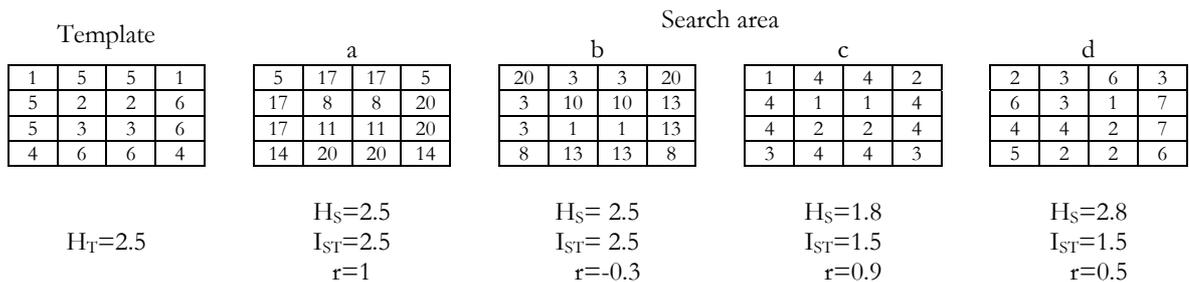


Fig. 1.13 Dependence of correlation coefficient r and mutual information I_{TS} on the relation between grey values of the template and the search area. H is entropy of an image patch. For explanation a – d see the text above.

The correlation coefficient is high if a linear dependence between image patches exists which is the case a) and c). Mutual information is high if an uncertainty about a template knowing a search area is low. In the case a) and b) there is a unique relation between each grey value from the template and search area. According to the formula 1.12 the joint probability is equal to marginal probabilities $p_{TS}(t,s)=p_T(t,s)=p_S(t,s)$. The information contained in the template and search area equals, i.e. $H_T(t,s)=H_S(t,s)=I_{TS}(t,s)$. As soon as the correspondence between grey values in the template and search area loses its uniqueness, i.e. the conditional entropy increases, the mutual information decreases as it can be seen in the case c) and d). Due to conditions under which the aerial images are taken and also due to their further processing (e.g. in case of the orthoimages), the image patches correspond more to case c) and d). When

scanning the search area with the template, correlation coefficient looks for a position fitting to a linear radiometric model. Mutual information seeks the position where the amount of information that the template contains about a section of the search area is maximal. Therefore the found position can differ significantly.

C3) Correlation coefficient, mutual information, and image distance

The results of matching 122 image patches containing road crosses obtained in the two previous tests were used again. The positions of road crosses were known in the first orthoimage from which the templates were derived and were therefore used as reference values. Tab. 1.11 summarises the results of comparison of positions of best fit derived by means of different similarity measures with reference data. Different radiometric corrections were again applied prior to calculating the image distance and mutual information. Results of the image distance with pre-processing of a search area by means of a linear function were not used because they did not differ from results obtained by means of correlation coefficient (see Tab. 1.9). A root mean square error of one pixel in each co-ordinate was expected from a comparison. Based on an assumption of the normal distribution of random errors, differences greater than three pixels were considered as outliers. It must be mentioned that outliers found in this test are not only results of mismatching but can also be caused by errors in orientation parameters or DTMs used for deriving orthoimages.

Similarity measure		RMSE		mean		Number of outliers
		dX [pel]	dY [pel]	dX [pel]	dY [pel]	
Correlation coefficient		3.2	3.7	-0.3	0.3	17% (21 points)
Image distance	No radiometric pre-processing	4.2	4.3	0.2	0.3	20% (24 points)
	Histogram stretch	3.7	4.3	0.0	0.2	21% (26 points)
Mutual information	No radiometric pre-processing	8.1	8.2	0.0	-0.2	63% (77 points)
	Histogram stretch	7.2	7.1	-0.1	-0.5	52% (63 points)
	Linear function	3.3	3.7	-0.1	0.3	20% (25 points)

Tab. 1.11 Comparison of positions of best fit of 122 road crossings obtained by means of different similarity measures with reference values.

The values in Tab. 1.11 show that the correlation coefficient and mutual information with radiometric adjustment of a search area by means of a linear function revealed almost the same results. Results obtained by means of image distance are of about 0.5 to 1 pel worse. Based on RMSE values and the number of blunders, it can be concluded that methods based on mutual information without any radiometric corrections or in a combination with a histogram stretch of a search area into a range of a template are not suitable for the tested images.

D) Combination of similarity measures for the detection of mismatches

The performance of the three similarity measures was studied further with respect to their accuracy, reliability, and possibility of their combination in order to decrease the number of mismatches.

D1) Position of the best fit calculated for each of similarity measures

The tests carried out until now have shown that after applying some radiometric corrections, all three similarity measures reveal almost the same results after comparison with the reference data. The amount of outliers is about 20% (see Tab. 1.11). The following tests should show whether there is any possibility to exclude or at least reduce the outliers by combining the measurements based on different similarity measures. The results obtained by means of correlation coefficient, mutual information in a combination with a linear function for radiometric corrections and image distance with no radiometric pre-processing will be used.

Assuming that the found position of the best fit should be the same regardless of applied similarity measure, a search was done to find out how many outliers remained when points with differences in positions obtained by means of different similarity measures were excluded. Because of the test material the differences up to one pixel were also considered as acceptable. The results are summarised in Tab. 1.12.

Combining two measures and allowing differences of one pixel decrease the number of outliers to approximately 30-40% of an original amount. At the same time the number of points included into the calculation also decreases down to about 75%. The criterion of zero differences in obtained positions is very sufficient for an elimination of outliers if results obtained by means of mutual information are involved. The cost is a high number of excluded points. If such a criterion is applied, the distribution of remaining points has to be watched carefully to assure that the whole area of interest is covered. Image distance and correlation

coefficient are from their definition close to each other and therefore a criterion of zero differences does not exclude as many points as in the case of the comparison with mutual information. All the outliers were excluded when the criterion of zero differences was applied for all three measures, i.e. positions obtained by means of three similarity measures had to be identical. Only 33% of points then remained in calculation, which is rather small amount.

Correlation coefficient (r) and mutual information (I)					
$dX_{r-I} = 0$ [pel] and $dY_{r-I} = 0$ [pel]			$dX_{r-I} \leq 1$ [pel] and $dY_{r-I} \leq 1$ [pel]		
Excluded points	Excluded outliers		Excluded points	Excluded outliers	
	r	I		r	I
69% (84 p.)	90% (19 p.)	92% (23 p.)	25% (30 p.)	67% (12 p.)	60% (15 p.)

Correlation coefficient (r) and image distance (D_N)					
$dX_{r-D} = 0$ [pel] and $dY_{r-D} = 0$ [pel]			$dX_{r-D} \leq 1$ [pel] and $dY_{r-D} \leq 1$ [pel]		
Excluded points	Excluded outliers		Excluded points	Excluded outliers	
	r	D_N		r	D_N
29% (35 p.)	57% (12 p.)	63% (15 p.)	11% (14 p.)	43% (9 p.)	50% (12 p.)

Image distance (D_N) and mutual information (I)					
$dX_{D-I} = 0$ [pel] and $dY_{D-I} = 0$ [pel]			$dX_{D-I} \leq 1$ [pel] and $dY_{D-I} \leq 1$ [pel]		
Excluded points	Excluded outliers		Excluded points	Excluded outliers	
	D_N	I		D_N	I
71% (87 p.)	92% (22 p.)	92% (23 p.)	27% (33 p.)	75% (18 p.)	68% (17 p.)

Correlation coefficient (r), image distance (D_N) and mutual information (I)							
$dX_{r-I} = 0$ [pel] and $dY_{r-I} = 0$ [pel]				$dX_{r-I} \leq 1$ [pel] and $dY_{r-I} \leq 1$ [pel]			
and				and			
$dX_{r-D} = 0$ [pel] and $dY_{r-D} = 0$ [pel]				$dX_{r-D} = 0$ [pel] and $dY_{r-D} = 0$ [pel]			
Excluded points	Excluded outliers			Excluded points	Excluded outliers		
	r	D_N	I		r	D_N	I
77% (94 p.)	100% (21p.)	100% (25p.)	100% (24p.)	44% (54 p.)	90% (19p.)	92% (22p.)	84% (21p.)

Tab. 1.12 Excluding outliers based on a comparison of positions of best fit obtained by means of a combination of different similarity measures

After evaluating all possible solutions, the following criteria were chosen as the most optimal:

- Differences between positions obtained by means of correlation coefficient and mutual information should not differ more than 1 pixel in both co-ordinates.
- Differences between positions obtained by means of correlation coefficient and image distance should be equal.

Based on these criteria, 44% of points were excluded. 1 % of original outliers (two points) remained in the data set obtained by means of correlation coefficient and image distance and

2% of original outliers (four points) were present in the data obtained by means of mutual information. In comparison to Tab. 1.11 the root mean square error calculated from remaining 56% of points dropped down to $RMSE_x=1.0$ pel, $RMSE_y=1.3$ pel. The performed test showed that it is possible to eliminate the amount of mismatches by comparing the positions of best fit acquired by means of different similarity measures. The disadvantages of the method are:

- The amount of ‘good’ points excluded is relatively high.
- Three independent searches increase calculation time significantly, especially if the template and search areas are large.

D2) Position of the best fit calculated by means of cross-correlation; image distance and mutual information are used as attributes

As Tab 1.11 shows and as mentioned above, the results obtained by means of correlation coefficient and mutual information revealed almost the same root mean squares errors after comparing with the reference values. Calculation of mutual information includes creating and evaluation of a 2D histogram. When least squares adjustment for derivation of a radiometric shift and scale of a section of a search area is included, the calculation of a position of the best fit takes much longer in comparison to the calculation of correlation coefficient. Therefore a new test was performed. The position of the best fit was calculated only by means of the correlation coefficient. The values of image distance and mutual information were calculated at that position too and added as attributes to the point. The goal was to find out whether it was possible to exclude mismatches by evaluating/thresholding these attributes.

First, the differences of the found position of the best fit and reference values were plotted together with obtained values of three studied similarity measures. It is obvious from Fig. 1.14 that the relation between the differences and similarity measures is not functional. Setting a threshold for a similarity measure reduces number of outliers but on the other hand it can also lead to excluding a large number of correct points. Based on Fig. 1.14 thresholds $r>0.8$, $D_N<16$ and $I>3$ were set up. The meaning of the thresholds is that only measurements where the similarity measure reaches the given threshold are accepted. The influence of the thresholds on the elimination of outliers and reducing the number of measurements is summed up in Tab. 1.13. A threshold for correlation coefficient seems to be the most effective regarding high amount of excluded outliers and at the same time relatively low number of excluded correct measurements (almost a half of excluded points are outliers).

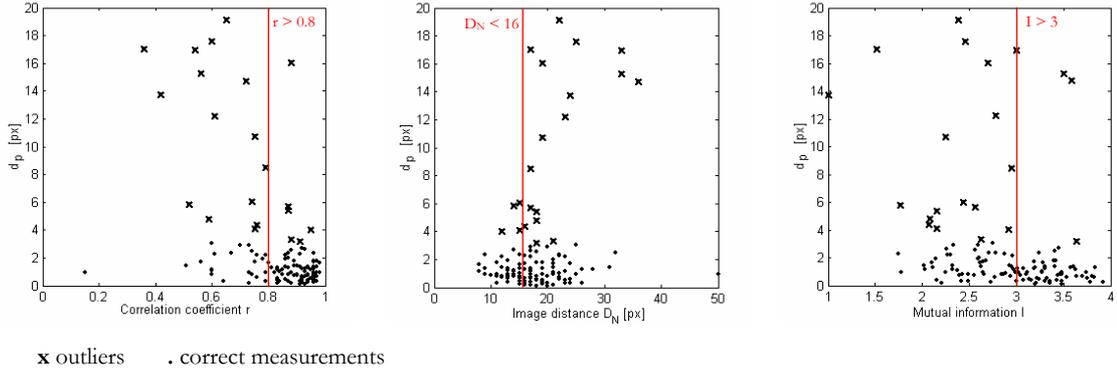


Fig. 1.14 Similarity measures and differences between positions of the best fit determined by means of correlation coefficient and reference values $d_p = (dX^2 + dY^2)^{1/2}$. A cross stands for outliers, a dot for a correct measurement.

Thresholds for similarity measures	Excluded points	Remaining outliers	Excluded outliers
Correlation coefficient $r > 0.8$	32% (39 points)	29% (6 points)	71% (15 points)
Image distance $D_N < 16$	65% (79 points)	19% (4 points)	81% (17 points)
Mutual Information $I > 3$	55% (67 points)	14% (3 points)	86% (18 points)

Tab. 1.13 Influence of setting thresholds for similarity measures on number of measurements and outliers excluded from calculation. The calculation was done for 122 points that included 21 outliers.

Second, thresholds for some other measures derived from mutual information and image distance were investigated. Among others entropy correlation coefficient $ECC(\mathbf{g}_T, \mathbf{g}_S) = 2I(\mathbf{g}_T, \mathbf{g}_S) / (H(\mathbf{g}_T) + H(\mathbf{g}_S))$ (Maes et al., 1997) and a ratio between image distance (D_N) and contrast (standard deviation) of the template and search area $\sigma_{TS} = ((\sigma_T^2 + \sigma_S^2) / 2)^{1/2}$. Similarly to Fig. 1.14, a relation between these two measures and found differences is shown in Fig. 1.15. While a threshold for the entropy correlation coefficient $ECC > 0.5$ excludes 63% (77 points) of all measurements and 86% (18 points) of the outliers, a threshold $D_N / \sigma_{TS} < 0.65$ gives a more favourable score 31% (41 points) of excluded measurements and 81% (17 points) of excluded outliers. Under assumption $\sigma_{TS} = \sigma_T = \sigma_S$ the threshold $D_N / \sigma_{TS} < 0.65$ corresponds to the threshold for correlation coefficient $r > 0.79$. Due to differences in contrast in the template and search areas the thresholds do not have the same effect.

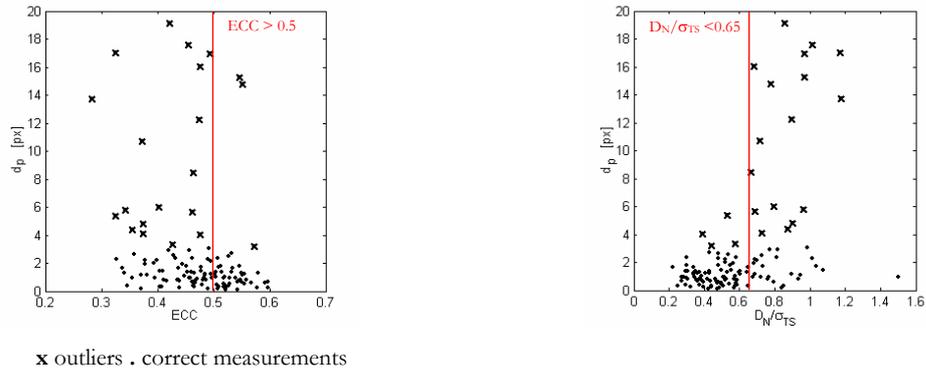


Fig. 1.15 Entropy correlation coefficient ECC and ratio between image distance and contrast (standard deviation of grey values) of a template and search area in relation to differences between positions of the best fit determined by means of correlation coefficient and reference values $d_p = (dX^2 + dY^2)^{1/2}$.

By comparing Fig. 1.14 and 1.15, it can be concluded that thresholding the measure D_N/σ_{TS} brings the best result. Root mean square error in differences of found positions to the reference values for measurements fulfilling a condition $D_N/\sigma_{TS} < 0.65$ is equal to $RMSE_X = 1.1$ pel, $RMSE_Y = 1.1$ pel. Fig. 1.16 shows the distribution of outliers, excluded points and correct points.

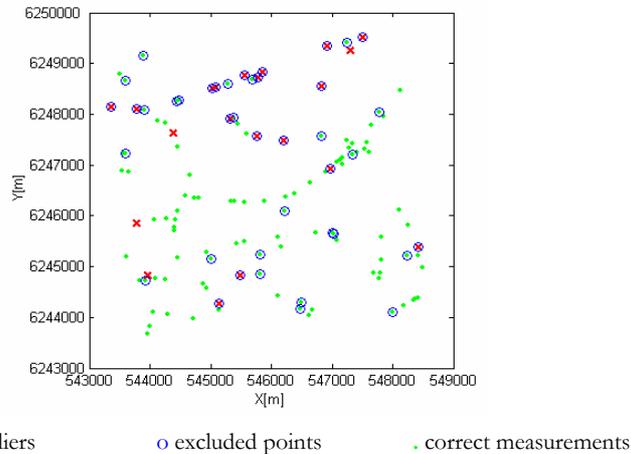


Fig. 1.16 Distribution of outliers (red crosses), excluded points (blue circles, $D_N/\sigma_{TS} \geq 0.65$) and correct measurements (green dots). X, Y are co-ordinates in the UTM co-ordinate system, zone 32.

In the last sections three similarity measures applicable for area based matching were studied, namely correlation coefficient, image distance and mutual information. The aim was to find out about their differences and suitability for photogrammetric applications. All three measures are based on the comparison of grey values in a template and in a search area. Depending on image acquisition and type of matched images, the radiometric differences

between image patches can be very large. Therefore a linear radiometric balancing by means of histogram stretch or linear parameters (scale and shift) derived by least squares adjustment (LSA) was included into calculations. In agreement with its definition, positions of best fit obtained by means of correlation coefficient were the same regardless of a linear radiometric adjustment. Image distance performs similarly to correlation coefficient. If linear parameters for balancing a search area to a template are derived by means of LSA, the results obtained by image distance are identical to those derived by correlation coefficient. Mutual information has not been investigated in the field of photogrammetry so far. The promising results were acquired in medicine applications with MR and CT images (Maes et al., 1997). In the performed tests, the results obtained by means of mutual information were acceptable and on the same level as from correlation coefficient only when radiometric adjustment of a search area by means of linear function was carried out. In comparison to correlation coefficient and image distance, the calculation of mutual information including derivation of linear radiometric parameters by means of LSA is time consuming and therefore not optimal for processing of large amount of data.

A lot of outliers can be introduced in area based matching. Therefore possibilities of their reduction by means of comparing results obtained by use of different similarity measures or setting thresholds for similarity measures were investigated. Performing image matching several times based on different similarity measures and comparing differences in found positions brings good results but is rather costly. Finding a position of the best fit by means of one measure e.g. correlation coefficient and calculating and evaluating other parameters at this position is a more applicable solution. The results showed that it is not possible to eliminate all outliers but their amount can be reduced significantly. Set thresholds differ from data set to data set depending on quality of images and especially objects to be matched.

The test material used in previous sections consisted of two orthoimages derived with a time difference of four years. It should be emphasised again that found differences in position are not necessarily caused only by errors in matching but also by errors/differences in orientation parameters and in the DTM. Therefore measurements that were excluded should be investigated further in order to find out the errors in the original data sets (see also chapters 2 and 3).

Based on stated conclusions, only the correlation coefficient, normalised image distance in combination with the standard deviation of grey values (contrast), and least squares matching

will be used in practical applications in the following chapters. Mutual information will not be investigated further.

1.2.2 Feature based methods

In contrary to the area based matching (ABM) that operates directly on grey values, **feature based methods** (FBM) are based on matching extracted features as points, edges, or regions. When measuring conjugate points in the image, the human operator detects and concentrates on abrupt changes in grey values (the image function) that often correspond to physical boundaries in object space rather than grey values themselves. When considering the human visual system as superior to automated systems, feature based methods are in better agreement with this system than area based methods (Schenk and Hofmann, 1986). Feature based matching procedures consist of three steps (adapted from Förstner, 1986):

- selecting distinct features (points, edges) in the images separately
- building up a preliminary list of candidate pairs of corresponding features based on a chosen similarity measure
- deriving a final list of feature pairs consistent with an object model

In the following text basic concepts of FBM are described. Most attention is paid to interest points and their combination with area based methods. This is an approach applied in most of professional photogrammetric software packages for an automatic measurement of tie points and derivation of elevation models.

1.2.2.1 Interest points

Area based matching performs best in areas of high contrast. Such points (or better small windows of the size e.g. $7 \times 7\text{pel}^2$) that can be described by high variances in grey values, high autocorrelation function and steep gradients are called interest points and they are derived by means of interest operators. According to (Förstner, 1986) the interest points should fulfil requirements of distinctness, invariance to expected geometric and radiometric distortions, stability (a point should appear in all matched images), seldomness (confusions in areas with repetitive patterns should be avoided), and interpretability. The procedure of finding interest points in each of matched images consists of two steps:

- calculating characteristic parameters for each window in the image or its selection
- comparison the values of these parameters with given thresholds.

The characteristic parameters are different for each operator, but basically depend on grey values (texture and structures) inside the evaluated window. Only windows which parameters'

values are greater/smaller than given thresholds are accepted as interest points. A list of interest points for each of matched images with their pixel co-ordinates (a centre of the evaluated window) and their description (parameters) is the result of this process. In (Luhmann and Altrogge, 1986) three **interest operators**, namely the operator of Moravec, Förstner, and Dreschler are described and evaluated. The results show that the Moravec and Förstner operators perform better with finding interest points under different geometric conditions.

The **Moravec operator** is based on assumption that an interest point has high variances in all directions, e.g. row, columns and both diagonals within an operator window $(2k+1) \times (2k+1)$ pixels. The interest point is accepted, if a characteristic parameter M (see formula 1.14) is greater than a given threshold.

$$M = \min \left\{ \begin{array}{l} \sum_{r=n-k, c=m-k}^{n+k, m+k} (g(r, c) - g(r, c+1))^2 \\ \sum_{r=n-k, c=m-k}^{n+k, m+k} (g(r, c) - g(r+1, c))^2 \\ \sum_{r=n-k, c=m-k}^{n+k, m+k} (g(r, c) - g(r+1, c+1))^2 \\ \sum_{r=n-k, c=m-k}^{n+k, m+k} (g(r, c+1) - g(r+1, c))^2 \end{array} \right. \quad (1.14)$$

k defines the size of the operator window

n, m pixel co-ordinates of the centre of the operator window within an image g

$g(r, c)$ grey value at the position (r, c)

The **Förstner operator** is connected to least squares matching method. The covariance matrix of the estimated shifts Σ can be written as following (formula 1.15):

$$\Sigma = \sigma_0^2 \begin{pmatrix} \sum g_R^2 & \sum g_R g_C \\ \sum g_R g_C & \sum g_C^2 \end{pmatrix} = \sigma_0^2 \begin{pmatrix} q_{RR} & q_{RC} \\ q_{RC} & q_{CC} \end{pmatrix} = \sigma_0^2 \mathbf{Q} = \sigma_0^2 \mathbf{N}^{-1} \quad (1.15)$$

σ_0 standard deviation of a unit weight

g_R, g_C ... gradients in row and column directions

\mathbf{Q} cofactor matrix

Without knowing the standard deviation σ_0 , a shape of an error ellipse (i.e. its semimajor and semiminor axes) can be derived as eigenvalues $\lambda = \{\lambda_1, \lambda_2\}$ of the cofactor matrix \mathbf{Q} . A good match has to fulfil two requirements:

- a) the error ellipse should be small, i.e. $\lambda \rightarrow \min$

The condition leads to finding a minimum of characteristic polynomial of \mathbf{Q} with the solution 1.16:

$$\begin{aligned} \lambda^2 - 2(q_{RR} + q_{CC})\lambda + q_{RR}q_{CC} - q_{RC}^2 &= 0 \\ \lambda \rightarrow \min \Leftrightarrow \lambda &= q_{RR} + q_{CC} = \text{tr}(\mathbf{Q}) = w \rightarrow \min \end{aligned} \quad (1.16)$$

The parameter $w = \text{tr}(\mathbf{Q})$ that characterises an interest point should be small. Thus, only interest points with $w < w_{\text{threshold}}$ are accepted. According to (Kraus, 1997) the third of the mean value of w calculated in an image or its large extract is suitable as a threshold value.

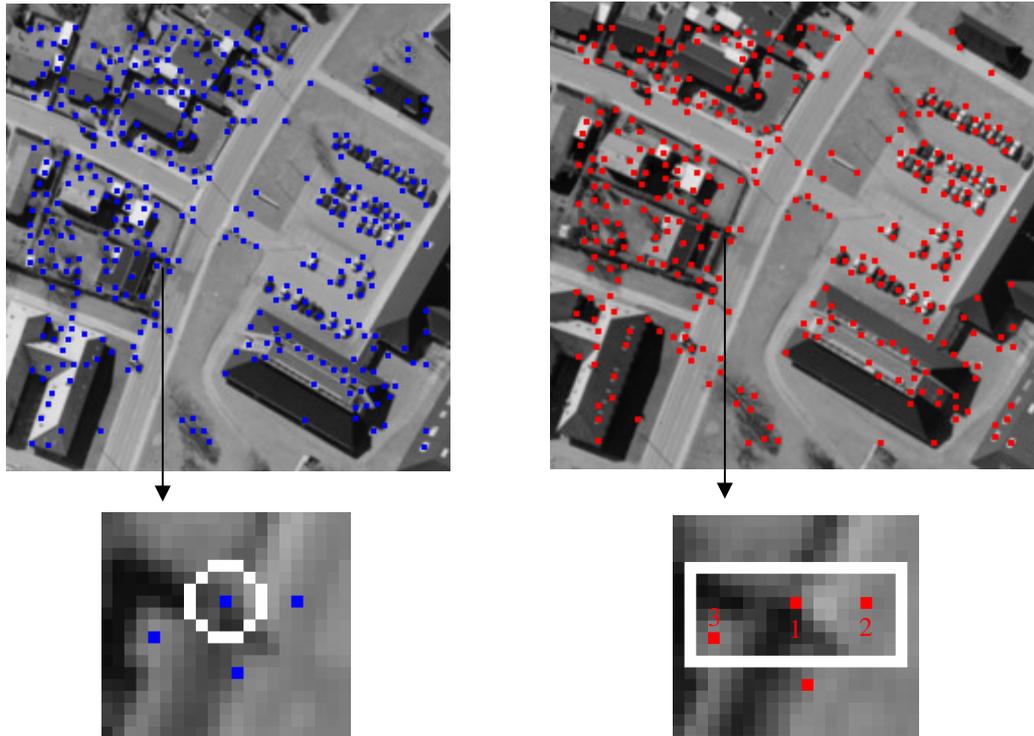
- b) the error ellipse should be close to the circle (otherwise the match is not well defined in one direction), i.e. $\lambda_1 = \lambda_2$.

This condition can be described by the equation 1.17:

$$1 - \frac{(\lambda_1^2 - \lambda_2^2)^2}{(\lambda_1^2 + \lambda_2^2)^2} = \frac{4\text{Det}(\mathbf{N})}{\text{tr}(\mathbf{N})^2} = q \quad (1.17)$$

Maximal value of the criterion q is equal to 1 that corresponds to uniform accuracy in all directions. According to (Kraus, 1997) the threshold for accepting interest points can be set as $q > 0.75$.

Both Moravec and Förstner operators allow for finding interest points with subpixel accuracy (Tang and Heipke, 1996 and Kraus, 1997). In case of the Moravec operator the centre of gravity with respect to the squared grey value differences in the window is calculated. In case of the Förstner operator discrete w values are approximated with a continuous function which minima is sought (similar to finding maximum of a 2nd order polynomial approximating a correlation function in Appendix B.1). In (Tang and Heipke, 1996) it is also pointed out that calculations according to the Moravec operator require less calculation time. Results in (Luhmann and Altrogge, 1986) show that Förstner operator gives better precision, is rotation invariant and is less affected by local contrast variations.



Left image of a stereopair with derived interest points a detail of a chosen interest point

Right image a stereopair with derived interest points and a detail of conjugate point candidates

	Correlation coefficient	Gradients	
		g_r	g_c
1	0.82	-58	80
2	-0.27	-5	-21
3	-0.12	2	15
Gradients of a chosen interest point (left image) $g_r = -35$ $g_s = 20$			

Fig. 1.18 Example of finding conjugate points in an image stereopair by means of feature based matching. An interest operator is first applied for both images in order to derive distinct points. The search for a corresponding point is restricted along an epipolar line. Correlation coefficient or gradients in grey values can be used as similarity measures. An interest point in the right image with the highest correlation score or the most similar gradient values is considered as a corresponding point to the interest point from the left image.

After selecting interest points in all images to be matched, corresponding points have to be found. The easiest solution to this problem is based on combining epipolar geometry and area based matching by means of correlation coefficient. For example, in case of relative orientation approximations of orientation parameters are used in order to find a position of a chosen interest point from the left image in the right image. The correlation coefficient is calculated for all interest points within a rectangle placed at that position and oriented along an epipolar line. The size of the rectangle depends on accuracy of the orientation parameters

and an elevation model. The point with the highest correlation score is considered as the conjugate point. It must fulfil two conditions. First, the correlation coefficient must be higher than a given threshold. Second, one and the same interest point may only occur in a single pair (Kraus, 1997). Fig.1.18 illustrates the procedure of finding conjugate points by means of correlation coefficient and epipolar geometry. Similar to the correlation coefficient, a gradient in row and column directions can be used as a similarity measure (Pedersen, 1999b). As evident from Fig. 1.18, the correlation coefficient brings more distinct results. Another alternative for evaluating similarity between interest points is least squares matching. The decision about correspondence between interest points from two images can be more complex as shown e.g. in (Förstner, 1986).

In spite of reaching a threshold for similarity measures, mismatches can appear. They have to be excluded by means of robust adjustment procedures that are incorporated into calculation of image orientation or DTM generation. Because of an epipolar line constraint, outliers in the direction perpendicular to epipolar line are eliminated. Applying a so-called parallax bound that is set according to height differences of the terrain can reduce mismatches in the direction of epipolar lines (see also chapter 3). Tests with relative orientation of real images showed that after removing outliers a standard deviation of image co-ordinates σ_0 of 1/3 or even 1/4 of pixel size can be achieved (Förstner, 1986, Tang and Heipke, 1996).

1.2.2.2 Edges and regions

Feature based matching with edges and regions as primitives is included into this thesis in order to complete an overview of image matching methods that have been used for automatic orientation of aerial images. They were not implemented into practical applications described in chapters 2 and 3 and therefore only basic terms, procedures, and experiences found in literature are briefly mentioned without detailed explanations.

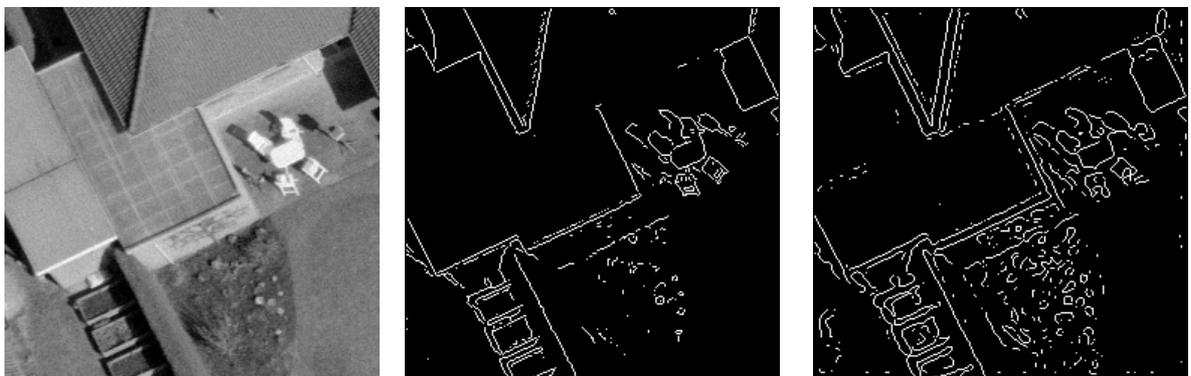
Edges can be describes as discontinuities in the grey level function e.g. grey values rapidly change within a small area. Edges usually correspond to boundaries of objects displayed in the image. The process of edge extraction is rather complex and consists of following steps (Schenk, 1999):

- detecting edge pixels, i.e. grey value discontinuities are detected by means of so called edge operators. A threshold for the differences in grey values has to be set in order to decide on an edge pixel.
- linking edge pixels to edges, i.e. connecting pixels belonging to the same edge

- grouping edges, i.e. identifying straight line segments, polylines, parallel lines, etc.

Edge operators detect grey value differences in the image. Some of them are based on calculating the first derivative of the grey level function – an extreme of the first derivative then locates an edge pixel. The **gradient operator** or **Sobel operator** can be named as examples. Both of them are direction dependent, one form of an operator detects horizontal edges and its transposed form vertical edges (see Fig. 1.19). The Sobel operator is less sensitive to noise due to including neighbouring pixels.

The **Laplacian operator** belongs to the group of the second derivative operators. It is direction independent. In order to suppress an influence of the noise in images, it is combined with the Gaussian operator that smooths the image. After applying the resulting **Laplacian of Gaussian (LoG) operator** on the original image, an edge pixel corresponds to the zero value in the convolved image (also referred as **zero crossings** e.g. transition from the positive to the negative convolution values or vice versa). The LoG operator as well as first derivative operators is described in detail in e.g. (Schenk, 1999, Russ, 1999). Fig. 1.19 shows examples of applying the Sobel and LoG operators on a section of an aerial image. The obtained binary images are the results of histogram thresholding (segmentation) after applying edge operators.



Original image

Edge detection

Sobbel operator		LoG operator				
horizontal edges	vertical edges	$\sigma = 2$				
1 2 1	1 0 -1	0.045	0.047	0.056	0.047	0.045
0 0 0	2 0 -2	0.047	0.317	0.715	0.317	0.047
-1 -2 -1	1 0 -1	0.056	0.715	-4.905	0.715	0.056
		0.047	0.317	0.715	0.317	0.047
		0.045	0.047	0.056	0.047	0.045

Fig. 1.19 Edge derivation by means of Sobbel and LoG operators. Size of the LoG operator is equal to $2\sigma+1$. Function 'edge' of MATLAB[®]'s Image processing toolbox was used for edge extraction and creating BW images.

Linking edge pixels into lines with a starting and end point or into regions is a rather complicated task. Principles of good continuation (extending a line to an edge pixel that causes the least change in orientation) or region growing (finding closed regions as zero crossing contours) can be applied (Schenk, 1999).

There are two different methods for matching derived edges:

a) Matching edge pixels

A chosen edge pixel is in e.g. the left image of a stereopair searched in the right image. A search is done along an epipolar line. Attributes as a sign in case of zero-crossings (indicates whether brighter/darker grey values are to left/right from the chosen edge pixel), orientation and strength (gradient across the edge), or a measured horizontal parallax are taken into account in order to find a corresponding pixel. The method is suitable for applications like DTM generation where orientation parameters of images are exactly known and therefore epipolar lines can be determined more precisely. For faster calculations normalised images are generated.

b) Matching entire edges

In comparison to the previous method, matching entire edges is suited also for images without a known orientation and can be therefore used for the relative orientation of images. The matching is based on finding an edge or a segment/segments of an edge within a search area e.g. in the right image that corresponds best to the shape of an edge chosen in the left image. For the purpose of 'shape' matching an edge and its segments are usually parameterised by the generalised **Hough transform** or represented by the **ψ - s curve** (Habib and Kelley, 2001, Schenk et al., 1991, Schenk, 1999).

Matching edges can be improved further by means of **relational (symbolic) matching** that considers relationships between edges (Schenk, 1999).

A lot of research has been done regarding matching edges and relational matching especially for the purpose of relative orientation. Nevertheless, the methods have not found many practical applications until now. Feature extraction and matching is often computationally expensive and requires number of parameters and thresholds that have must be chosen a priori (Heipke, 1996). Matching interest points is the technique that has been widely implemented into photogrammetric software packages. In higher levels of the image pyramid it is used in combination with correlation coefficient or gradient for improving approximations of orientation parameters while in the level with the finest geometric

resolution it is combined with least squares matching in order to achieve a high accuracy. To ensure reliability of results, the hierarchical approach is combined with epipolar geometry and highly redundant set of measurements.

1.3 Outliers, their detection and removal

All automated processes in photogrammetry based on image matching like relative orientation, aerotriangulation, and DTM derivation have to deal with erroneous measurements, i.e. mismatches. A combination of different similarity measures and matching techniques together with thresholding and applying additional constraints reduce their amount (see chapter 1.2.1.5, test D). The best matches within a local area are found but there is no guarantee that the measurements are consistent with respect to the chosen mathematical model. That means that if the number of measurements (n) exceeds the number necessary to uniquely determine the model (n_0), different combinations of n_0 measurements give different estimates of model parameters. Adjustment techniques are used to eliminate the inconsistency in the measurement. In general, three types of errors can appear in connection with measurement: random errors, systematic errors, and outliers. **Random errors** gain random size and magnitude under the same conditions of measurement. In case of absence **of systematic errors** they have the normal (Gaussian) distribution $N(\mu, \sigma^2)$ with the mean value $\mu=0$ and the standard deviation σ . In photogrammetric applications systematic errors connected to e.g. a sensor, atmosphere etc. can be modeled and corrected. **Outliers** are those errors that do not show a systematic character and do not fit to the considered error distribution.

An assumption of the normal distribution fits well with the adjustment method of least squares that minimizes the sum of squares of residuals (or corrections) between adjusted and original values. Because of its simplicity, calculation speed and also due to tradition, least squares adjustment (LSA or L_2 -norm) is the most applied method. A detailed explanation of LSA, its computational algorithms and post-adjustment analyses can be found e.g. in (Mikhail and Gracie, 1981). The real data sets seldom fulfill the demand for the normal error distribution. A 'longer tailed' distribution often appears, i.e. errors large in magnitude occur more frequent than expected in the normal distribution. These outliers (also gross errors or blunders) cause deviations from the normal distribution and can lead to a poor or erroneous estimation of the model parameters. Therefore they require a special care (Kubik et al., 1984, Rousseeuw and Leroy, 1987).

There are basically two ways how to deal with outliers – **data snooping** and **robust parameter estimation**. The first method uses LSA and then excludes the largest error based on the statistical tests. The process is repeated until all residuals fulfill criteria of the normal distribution. The second method minimizes a function of residuals that is different from the one used in LSA and is able to reject/recognize all or at least most of present outliers in one procedure (if they are in minority). The following chapters give a general overview of both approaches. The intention is only to explain the basic concepts and procedures that were applied in practical tasks in chapters 2 and 3 and not to give a detailed comparative study of methods, as it is beyond the concern of this thesis.

1.3.1 Data snooping

Data snooping is based on an assumption that there is only a single outlier in the measurements (observations) and that the linearization of a possibly non-linear problem has no significant effect (Kraus, 1997). The procedure consists of following steps:

- least squares adjustment
- calculating standardized residuals
- statistical test on standardized residuals

Standardized residuals can be derived in two ways – by dividing residuals by a known a priori standard deviation or by standard deviations of residuals derived after least squares adjustment. In the first case the standardized residuals are tested for the normal distribution, in the second case for the Student (t) distribution. The standardized residual with the highest magnitude that exceeds the critical value for redundancy $r=n-n_0$ and a confidence interval α is excluded from calculation as a outlier. The whole process is repeated until no further outliers are located (Kubik et al., 1984, Schwarz and Kok, 1993). The above described method was applied in the calculation of spatial resection and the comparison of orthoimages based on image matching in chapter 2.

The method has some disadvantages that restrict its wider application. First, only one outlier is detected within one computational cycle. A wrong decision about an exclusion of the observation increases probability of outliers in the data set. Second, outliers can remain unrecognized. Examples in (Rousseeuw and Leroy, 1987) demonstrate that a largest residual does not necessarily mean a blunder in the data set. In order to improve the method, Baarda introduced the concept of reliability (Kubik et al., 1984, Kraus, 1997). Such an extension of an original simple test theory with more parameters makes its application in automated processes

more complicated and the use of the method in photogrammetry is more or less limited to research.

1.3.2 Robust adjustment

Robust estimation methods can be characterized by low sensitivity to small deviations from the assumed error distribution, namely outliers (Huber, 1981). The sensitivity against outliers can be evaluated by two measures – **statistical (or relative) efficiency** and **breakdown point**. Statistical efficiency is the ratio between the lowest achievable variance of the estimated parameters and the variance obtained by the robust method. Breakdown point is the largest percentage of outliers that can be tolerated by the method before the breakdown occurs e.g. before the output estimate is arbitrary wrong. It is difficult to find a method which balances both parameters well. Least squares estimation can be used as an example of the most statistically efficient estimator but only in case of the normal error distribution. Already one outlier can cause a completely wrong solution, its break point is $1/n$ (which tends to 0% for an increasing number of measurements n) (Shahrabi, 2000). Evaluation of different robust estimators regarding breakdown points and other characteristics can be found in (Rousseeuw and Leroy, 1987).

L_1 -norm estimator that minimizes the sum of absolute residuals $\sum |v_i| \rightarrow \min$ is in its principle a simple robust adjustment technique. But due to computational difficulties the method did not find large application until the simplex algorithm that gives a solution to a linear programming problem was introduced. An example of a possible application of L_1 -norm method in automatic measuring procedures in photogrammetry is given in (Calitz and Rütther, 1994) who solve least squares matching as ‘least absolute’ matching. They concluded that the L_1 -norm method locates the matched point more accurately if the search patch contains occlusions or objects (pixels) that are not present in the template. The disadvantage of the method is a lower accuracy of the estimated parameters. Therefore (Gao et al, 1992) suggests two steps approach – first, finding and removing outliers through the robust testing procedure and second, processing ‘clean’ data by LSA to get a minimum variance solution. In statistical literature robust estimators have been widely studied with respect to the linear regression (Rousseeuw and Leroy, 1987). Fig. 1.21 shows performance of L_1 -norm in comparison to other adjustment methods.

Maximum likelihood or M-estimators represent another group of the robust estimators (Huber, 1981, Rousseeuw and Leroy, 1987). Their principle is in minimizing a function of residuals $\rho(v_i)$ that is symmetric with a unique minimum in 0 (formula 1.18).

$$\sum_{i=1}^n \rho(v_i) \rightarrow \min \quad (1.18)$$

In practical applications, the numerical solution of adjustment by means of M-estimators is iterative and equal to the method of weighted least squares ($\sum p_i v_i^2 \rightarrow \min$). In the first iteration a priori given weights are used. In the next iterations the weights of observations change according to the magnitude of residuals. The calculation continues until convergence is achieved (according to (Kubik et al., 1984) 3-20 iterations are necessary). An example of rather complex weighting function was suggested e.g. by Hampel and it can be found in (Jørgensen et al., 1984, Shahrabi, 2000).

A relatively extensive research regarding the robust adjustment was done in the Geodetic Institute of Denmark (today the National Survey and Cadastre) and Aalborg University (Krarup et al., 1980, Juhl, 1984, Kubik et al., 1984, Jørgensen et al., 1984). It resulted in the so called **Danish method** that describes the weight function by a formula 1.19.

$$p = 1 \quad \text{for } \text{abs}(v_i) \leq 2\sigma \quad (1.19)$$

$$p \text{ is proportional to } \exp(-c v_i^2) \quad \text{for } \text{abs}(v_i) > 2\sigma$$

p ... weight

v ... residual

σ ... standard deviation of measurement

c ... constant

In its principle the method searches for the largest number of mutually consistent measurements and uses only these measurements in least squares adjustment for determining unknown parameters. For the purpose of bundle adjustment a variant of an original method given by a formula 1.20 proved to be more efficient (Krarup et al., 1980).

$$1^{\text{st}} \text{ iteration } p = 1 \quad (1.20)$$

$$2^{\text{nd}} \text{ and } 3^{\text{rd}} \text{ iteration } p = \exp(-0.05 | v_i / \sigma |^{4.4})$$

$$\text{following iterations } p = \exp(-0.05 | v_i / \sigma |^{3.0})$$

Figure 1.20 shows the weight function for the second and fourth iteration.

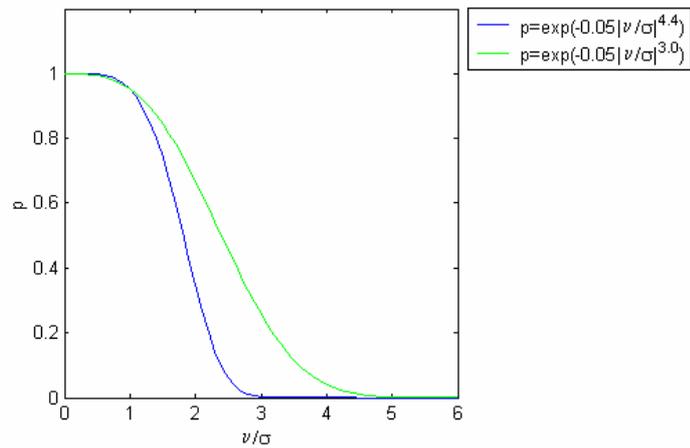


Fig. 1.20 Weight functions according to the formula 1.20. The first function (the blue curve) is rather tough and weights down all points where $|v/\sigma| > 2.5$. During the fourth and following iterations those points that are not outliers return to adjustment (the green curve).

The weight reduction function is empirical. It cannot be classified within an existing maximum likelihood theory of robust estimation (Juhl, 1984, Jørgensen et al., 1984). Nevertheless, the Danish method proved to be very effective in dealing with erroneous data (see Fig. 1.21). It also has favorable properties with respect to speed of computation and an easy implementation into programs designed for least squares adjustment. The standard deviation of measurements should be known a priori that is usually not a problem in photogrammetric orientation applications. Otherwise estimations of standard deviations calculated after latest iteration step have to be used. Some modifications of the method can be found in literature e.g. (Werner, 1984) or (Schickler, 1992).

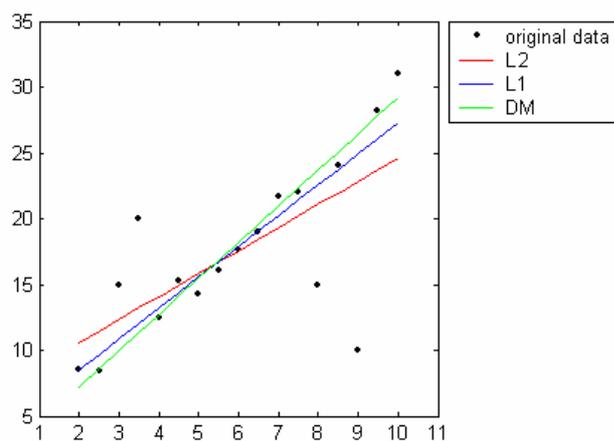


Fig. 1.21 Comparison of results of linear regression by means of least squares adjustment (L2), least absolute sum adjustment (L1), and the Danish method (DM). The original data contain four outliers for $x=3, 3.5, 8, 9$.

Random sampling is another method that has found an application in photogrammetry (Fischler and Bolles, 1981, Lacey et al., 2002, Shahrabi, 2000, Läbe, 1999). In contrary to other adjustment methods that use as much data as possible to obtain an initial solution and then try to discover outliers, according to the original idea of (Fischler and Bolles, 1981) **the random sample consensus** algorithm (RANSAC) constructs the solution from the minimum subset of data necessary for solving unknown parameters and enlarges this set with consistent data when possible. The process is repeated enough times for different randomly chosen subsets to ensure that there is e.g. a 95% chance that one of the subsets will contain only good data points. A solution that maximizes the number of points whose error measure is within a given threshold is sought. There are three parameters that have to be specified before the algorithm starts – the error tolerance used to determine whether a point is compatible with the estimated model, the number of subsets to try, and the threshold specifying the number of points used to imply that the correct model has been found. A detailed description of all three parameters can be found in (Fischler and Bolles, 1981).

The advantage of the RANSAC method is that it can evaluate any estimate of a set of parameters regardless of accuracy of a method used for derivation of those parameters. Results in (Shahrabi, 2000) with fitting 3D plane polygons over an existing digital surface model showed that random sampling techniques provided a good initial guess of the plane parameters, although all outliers were not excluded. Therefore, two stage robust estimation was suggested. First, RANSAC that has a high breakdown point (above 50%) provides a first guess of unknown parameters and if necessary also a standard deviation. In the second stage the M-estimator that yields good results only with good starting approximations and the number of outliers not exceeding 35% is applied. The recommendation of combining or applying different robust techniques on the same data in order to evaluate and possibly increase the reliability of the result can also be found in other literature (Kraus, 1997, Jørgensen et al., 1984).

Many other robust estimators can be found in literature (a detailed theoretical overview is given e.g. by (Huber, 1981)). In this chapter, an attention was paid only to four methods that were found as applicable for the automated processes in photogrammetry. Because of its simplicity in implementation and good empirical results regarding bundle adjustment, a modification of the Danish method according to (Juhl, 1984) was included into a calculation of spatial resection (see chapter 2 and Appendix A).

2. Automated exterior orientation of aerial images based on existing data sets

Determining parameters of interior and exterior orientation is the first task of photogrammetry that has to be solved to determine object co-ordinates of points measured in the images. Some aspects of interior orientation were discussed in the chapter 1.1. The parameters of exterior orientation can be basically determined in two ways:

- a) **Indirect orientation** that requires a measurement of control features (points, lines) both in object and image space. Depending on the number of images involved, it can be solved by following procedures:
 - Relative and absolute orientation of a stereopair (two images)
 - Aerotriangulation (n images)
 - Spatial resection (one image)
- b) **Direct orientation** by means of GPS and an inertial measurement unit (IMU). The measurement of control points is not basically needed. But it is used for a system calibration and when high accuracy is demanded.

The methods differ in accuracy and degree of automation that have been achieved.

Twelve orientation parameters of two overlapping images, a stereopair, are traditionally solved in a two steps procedure, namely relative and absolute orientations. **Relative orientation** (RO) solves five of unknown parameters by means of coplanarity or collinearity equations (Schenk, 1999) and its result is a 3D, vertical parallax free stereoscopic model. In the classical approach, at least five corresponding points (so called tie points) must be measured in the overlapping area in both images. Usually six well distributed tie points corresponding to O. von Gruber's model that was originally designed for analogue instruments are measured (Kraus, 2000). An extensive research of automatic methods of RO was carried out in late 1980ies and continued in 1990s. General solutions independent from approximations of both scale and rotations of images as relational matching or rotation invariant feature based matching have been developed. But due to complicated algorithms and time consuming calculations only a few of them have been implemented into modules for automatic RO that are nowadays standard parts of photogrammetric software packages. In case of aerial images, the scale factors and rotations do not differ very much within a stereopair or a strip of images and are approximately known. This allows for matching interest points along approximate epipolar lines that is together with a hierarchical approach (an image pyramid) and a robust

adjustment in common to most software modules for automatic RO. An accuracy between 0.2 and 0.4 pixel can be achieved (Heipke, 1997). In comparison to manual measurement, the automated approach uses much more tie points (up to one hundred or more per stereopair). However, tie points must be well distributed over the overlapping area. **Absolute orientation** (AO) solves the seven remaining unknown parameters. It is represented by a 3D conform transformation from the model space to object space co-ordinate system. Ground control points (GCPs) whose co-ordinates are measured both in the object space and in the images are used as identical points for this transformation. Automated measurement of control points is a rather complex task and a general solution has not been found yet. Nevertheless, suggestions for solving this task in specific applications have been done and will be discussed in detail in the section 2.1.

Aerotriangulation gives the solution to the orientation parameters of a block of overlapping images. In general, it can be calculated in two ways – by means of independent models and as bundle block adjustment (Kraus, 2000). In digital photogrammetry the second approach that is based directly on the collinearity equations is used. The measurement of tie points connecting images within a block can be fully automated by means of multi-image feature or area based matching (Ackermann, 1996a). Measuring of control points is done manually in most of the practical applications. Possibilities of automation of the process are the subject of the section 2.1.

Spatial resection can be considered as a special case of bundle adjustment. Only six orientation parameters of a single image are to be solved. If more than three GCPs are known, the orientation parameters are found by adjustment procedure (see Appendix A).

Differential GPS together with an inertial measurement unit that consists of gyroscopes and accelerometers allow for a direct determination of both the position of the projective centre and of the rotation parameters of exterior orientation. The procedure is often called **direct sensor orientation** and has been employed successfully e.g. in laser scanning (Wehr and Lohr, 1999). A lot of attention has been paid to the development of this promising technology during the last two decades. The test carried out by the European organisation for experimental photogrammetric research (OEEPE, now called EuroSDR) showed that accuracy of 5 to 10 cm in planimetry and 10 to 15 cm in height can be achieved by means of a direct camera (sensor) orientation for large scale images 1:5000. That is about 2 - 3 times worse compared to standard photogrammetric results, e.g. to bundle block adjustment.

Especially with respect of reducing vertical parallaxes that can appear in the individual stereo models, an improvement can be accomplished by an additional automatic measurement of tie points. Systematic errors in the GPS/IMU measurements and changes in system calibration parameters cannot be detected without including ground control points into the calculation. The method that combines GPS/IMU measurements with measurements of tie and optionally also control points in one adjustment is referred as **integrated sensor orientation** (Heipke et al., 2002). Direct sensor orientation will probably become a dominating technology for sensor orientation. Due to relatively high costs that have an influence on its limited implementation and the need of integrated sensor orientation whenever a higher accuracy is demanded, the measurement of tie points and control points will still be in use and the requirement for its automation will still be relevant in near future.

2.1 Automated measurement of ground control points

The goal of absolute orientation or aerotriangulation is to establish transformation parameters between the model or image co-ordinate system and the object co-ordinate system. In order to compute those parameters, information common to both systems must be known. This is accomplished through ground control points, e.g. well-defined points known in object co-ordinate system that are identified and measured in images. There are several requirements for the control points. They should be easy to identify, unique, and it should be possible to locate them with a high precision. Depending on their nature, two groups of ground control points can be distinguished – signalised points and non-signalised or natural points related to topographic features. While a spatial position of signalised ground control points has to be measured in the field, the position of non-signalised features can be determined from existing data sets as maps, orthoimages, or databases designed for this purpose. In the following paragraphs both kinds of ground control points are described together with possibilities for automation of their measurement. Nevertheless, higher attention is paid to the measurement of natural control points that is connected to the practical application of the determination of orientation parameters based on information from existing data sets described in chapter 2.3.

Three or four legged crosses are examples of targets used for signalling control points. They must be of appropriate size and shape with respect to the image scale. A good contrast between the target and the background is also of a high importance (Kraus, 2000, Schenk, 1999). In order to automate the process of control point measurement, area based methods of image matching can be applied. Templates appropriate to the shape of the target, image scale, foreground/background contrast etc. have to be created. In case of using rotation dependent

matching methods such as cross-correlation, a series of templates corresponding to different rotations must be produced. However, due to distortions caused by inclined surfaces that are not considered in the template derivation and difficulties of modelling contrast differences, the templates can vary from the actual images of targets significantly. It can result in low score for similarity measures and mismatches. Another requirement necessary to ensure good results are relatively accurate approximations of orientation parameters that allow for a determination of approximate positions of control points in the images.

Results of a research in automating the process of measuring signalised control points for the aerotriangulation purpose are given in (Gülch, 1995, Hahn, 1997). Gülch suggests a feature based approach based on region segmentation, contour extraction and shape analyses instead of matching an image of a control point with a template. The method assumes that the control point (target) appears as a quite homogenous region in the image and that it can be described by a closed boundary. Internal checks of the size, colour, and shape of the target as well as thresholds for the homogeneity criteria and parameters for contour extractions are included into the procedure. The test of the method carried out with a stereopair of colour images at the scale 1:4000 and a pixel size of 15 μm showed that regarding accuracy and reliability the proposed method is comparable to methods based on least squares matching of derived templates and images of the targets. In Hahn's approach, a library of templates corresponding to different types of control point targets is created. Multi-image least squares matching is carried out in order to provide consistent approximate locations of control points in all images (Hahn, 1996). An accurate location of a control point in images is obtained by weighted least squares matching with a template from the library. Weights are applied in order to decrease an influence of inhomogeneous background. The tests were carried out for a block of 36 digital images at the scale of 1:13 000 and the pixel size of 15 μm . 200 signalised targets of different shape were used for the orientation of the block. The best result were achieved for squared targets regarding both accuracy of aerotriangulation ($\sigma_0=2.8\mu\text{m}\approx 0.2\text{px}$) and the success rate (99.4%).

The presented methods showed quite some potential for automation of measurement of signalized control points. Nevertheless, the signaling and maintaining control points for mapping purposes is rather costly and therefore methods using natural control points have been investigated. When using natural objects for the orientation the images, the term 'control feature' describes the situation better than 'control point' because not only point like objects

e.g. manhole covers or drain gratings but also line objects such as roads, area type objects as road intersections or 3D objects as buildings can be used as control information. The object co-ordinates of control features can be determined by terrestrial measurement (e.g. by GPS) or taken from suitable data sets. Nowadays, in many countries both technical maps of the cities and topographic maps exist in a digital form as 2D or 3D topologically structured vector maps with all the data organized in a database. Together with orthoimages and digital terrain models (DTMs) that provide height information if necessary, they are parts of nation-wide or local geographic information systems (GIS). Finding a position of control features can easily be solved by proper questioning the geographic database. However, the maps and databases have to be kept up to date in order to separate the orientation and map revision tasks.

If a data set of a suitable accuracy already exists the process of the control point measurement and the determination of parameters of exterior orientation can be automated, fully or to a high degree. The procedure generally consists of following steps (adapted from (Heipke, 1997)):

- Selecting control information, its extraction from an appropriate data set
- Defining primitives to be extracted, taking into account the appearance of the control information in the images
- Extracting the primitives from the image
- Matching the primitives with the control information
- Computing parameters of exterior orientation

The type of objects used for exterior orientation and therefore the choice of an appropriate data set used for an extraction of these objects depends on the image content, the scale and geometric resolution of images as well as on the requirement for accuracy that should be achieved. Control information and the primitives derived from images can have a vector character (straight lines, polygons etc.) and they are therefore suitable for feature based or relational matching. Another possibility is their representation in the form of image patches when the correspondence is found by means of area based matching. In the following chapter several examples of exterior orientation of images based on existing data sets are described.

2.2 Overview of methods for automatic exterior orientation of aerial images

The investigation of new methods for exterior orientation of aerial images was the goal of the **OEEPE test 'Automatic Orientation of Aerial Images on Database Information'** (Höhle, 1999a). The new methods were based on using existing data sets and were evaluated with regard to accuracy, speed and the level of automation. The methods were supposed to be efficient, fast, and low-cost in comparison with a classic aerotriangulation approach. The demands on digital maps, orthoimages or other data to be stored for orientation of new images should have been also defined. The test of the OEEPE project included two tasks. **Task A** was designed for large-scale images of urban areas. An image stereopair at a scale 1:5000 and a pixel size of 15 μm and a technical map with an accuracy of $\sigma_p=0.1 - 0.2$ m in planimetry and $\sigma_z=0.15 - 0.3$ m in height were available. Absolute orientation of a medium scale image using an existing orthoimage and a height model was the aim of **Task B**. A new aerial image with the scale of 1:27 000 and geometric resolution 30 μm , an existing orthoimage generated from images 1:25 000 with ground sampling distance of 0.8m, and a digital surface model (DSM) with accuracy about 0.5 m were provided as a test material. The time difference between images was two years. Seven different automatic or semiautomatic methods were developed and applied to the test data. The two following chapters give the summary of used procedures.

2.2.1 Methods based on information from a digital vector map

Three dimensional digital vector maps of a different accuracy and a different level of detail have been created and maintained in many countries. According to national standards, accuracy at cm level and high degree of detail is required for technical maps of cities as well as road, water supply, or sewage maps. In an ideal case, all the mentioned data are saved in one geographic information system. The extraction of well defined point-like objects as manhole covers or drain gratings, as well as line objects as edges of roads and gable lines of roofs can be carried out fully automatically. Further use of extracted information depends on matching strategy for image orientation.

A software package solving the exterior orientation by means of linear features derived from the parcel map has been developed in the Finnish Geodetic Institute. The exterior orientation is solved by the coplanarity equations. The authors emphasize a possibility of choosing weights for the least squares adjustment according to the line length, line definition and type

of control feature (based on GIS information, e.g. a ditch is worse defined than a roof line). For the test purpose the lines (edges) in the image were measured manually but this step could be automated (Karjalainen and Kuittinen, 1999).

The method described in (Jedryczka, 1999) uses the Laplacian operator for extracting edges in both aerial images of a stereopair. Using approximations of orientation parameters the lines from a vector map are projected into images. Edges close to projected lines are found in both black and white images of a stereopair obtained by the edge operator. Edge pixels are matched by means of image distance or correlation coefficient. The set of candidates is corrected by calculating relative orientation. The calculation of the parameters of exterior orientation follows. The process is repeated for each level of image pyramid. In the level with highest geometric resolution several iterations might be necessary in order to achieve the required accuracy.

The system for Automatic Model-based Orientation (AMOR) developed at University of Bonn is also based on edge detection and matching. For image orientation it uses 'control point models', e.g. 3D wireframe models of houses (Schickler, 1994, Läbe and Ellenbeck, 1996). Using approximate orientation parameters the control point models are projected into an image. In a close surrounding of a projected control point straight line segments are extracted. Matching between image edges follows. The goal of the next step is to eliminate outliers, e.g. incorrectly located control point models. The formerly used RANSAC method was exchanged to more robust complete search. Orientation parameters are then calculated by spatial resection with image edges as observations and with robust adjustment (weighted least squares). The last step is a self-diagnosis that evaluates the result of robust estimation with respect to precision and geometric configuration. The system has been integrated into the photogrammetric software package SOCET SET (Leica-Helava) and successfully used for orientation of aerial images for orthoimage production in the Survey department of North-Rhine Westphalia. A database of control point models has been created and maintained for that purpose. Approximations of orientation parameters are obtained by manual measurement on a digitizer (Läbe, 1999 and 2003).

The method presented in (Pedersen, 1996 and 1999a) does not extract features in the images as all of the above mentioned approaches but converts a vector presentation of map features into a raster form. The derived templates are matched by means of cross-correlation with the image and the position of the best fit is found. For obtaining subpixel accuracy an

approximation of the correlation function by a 2nd order polynomial is applied (see Appendix B.1). The strategy of the solution is based on the principle of redundancy and hierarchy. Hierarchy is not used only in a sense of an image pyramid but also of an object pyramid – road crossings are used for the orientation of upper image pyramid levels while manhole covers and drain gratings for the orientation of the level with highest geometric resolution. Orientation parameters of a stereopair are derived by means of robust bundle adjustment. Results from the higher pyramid level are used as approximations in the lower level.

The accuracy of the methods was evaluated by measurement of 25 checkpoints. Most of the methods result in root mean square error in X and Y co-ordinates between 7 cm and 10 cm (1 or 1.3 pel) and 13 cm in height (0.017% of the flying height). The obtained results well correspond to the accuracy of the technical map. Only methods of Läbe and Pedersen are quoted as fully automatic. Nevertheless, the potential of a full automation of the two other methods is obvious (Höhle 1999a).

2.2.2 Methods based on existing orthoimages and DTM

During the last years an orthoimage production increased considerably, a nation-wide orthoimage coverage is a standard map product. The orthoimages are updated with an average period of two to five years. Therefore alternative methods for orientation of aerial images with accuracy sufficient for an orthoimage production that would supplement or replace traditional aerotriangulation are required. After the first generation of orthoimages and the digital terrain model are produced, they can provide control information for a derivation of parameters of exterior orientation of a new aerial image. The condition is that both old orthoimage and DTM are free from outliers and systematic errors and their accuracy corresponds to national standards for orthoimage production. Description of three methods for image orientation based on mentioned data sets follows. The last method that matches image patches containing well-defined objects derived from an orthoimage and a new aerial image is further developed and tested in chapters 2.3 and 2.4. As mentioned at the beginning of the chapter 2.2, a DSM was used for the OEEPE test, Task B.

A novel computational method for deriving exterior orientation parameters is given by (Paszotta, 1999). A similarity measure between a new image and an existing orthoimage is described as a function of orientation parameters. Thus, the set of orientation parameters for which the similarity measure reaches its extreme is considered as a final solution. Image

distance is used as similarity measure. In order to find its minimum, a gradient method or the stochastic Monte Carlo method is applied (Paszotta, 2000).

The method proposed in (Shan, 1999) is based on matching a high number (according to an author up to thousands) of interest points derived in nine evenly distributed patches both in the orthoimage and a new aerial image. The correspondence between interest points is found by means of correlation coefficient and least squares matching. Additional topologic and geometric constrains are applied in order to avoid false matching. The calculation of orientation parameters is done by bundle adjustment with robust estimation of parameters.

The method presented in (Höhle, 1999b) is based on matching image patches derived from an orthoimage and a new aerial image by means of cross-correlation. Due to the time interval between taking the images, only the image patches containing time stable and well-defined objects are chosen. In order to eliminate erroneous matches, a threshold for the correlation coefficient is applied. After finding corresponding patches between the orthoimage and an aerial image, orientation parameters are calculated by means of spatial resection with robust adjustment.

All the methods were mostly solved by own developed software packages and reached a high level of automation in terms of matching, detection of outliers and calculation of orientation parameters. Regarding accuracy, the results of the methods differ in average at the level of 1-2 pel. Equally to most of methods based on a digital vector map, the main problem is the need of good approximations of orientation parameters. It can be solved by implementing hierarchical approaches (image and object pyramids) or by application of multi-antenna GPS or GPS-IMU technology.

2.3 Orientation of a single image based on existing orthoimage and DTM

A lot of countries produce orthoimages with the ground sample distance of 0.4 m, ..., 0.8 m from the photo scale about 1:25 000 especially for covering open land areas. The geometric resolution is suitable for planning or updating 2D digital topographic maps with content corresponding to maps of the scale 1:10 000 and smaller. In order to get orientation parameters of new images quickly and with low costs, methods based on existing orthoimages and DTM were suggested (chapter 2.2.2). The method described in (Höhle, 1999b) uses traditional and known approaches based on well-defined control information and area based matching. In comparison to the method developed by Shan (Shan, 1999), the user decides which areas (objects) are used for image matching. In this way both the number of mismatches and calculations can be reduced. Because this solution of image orientation was found useful and interesting from the practical point of view, it has become an object of further studies. Possibilities of an improvement of the method regarding both a level of automation and accuracy have been investigated. An influence of obtained orientation parameters on a new orthoimage derivation has been studied as well. Fig. 2.1 shows the placement of automatic orientation in the procedure of producing the next generation of orthoimages together with all necessary input data.

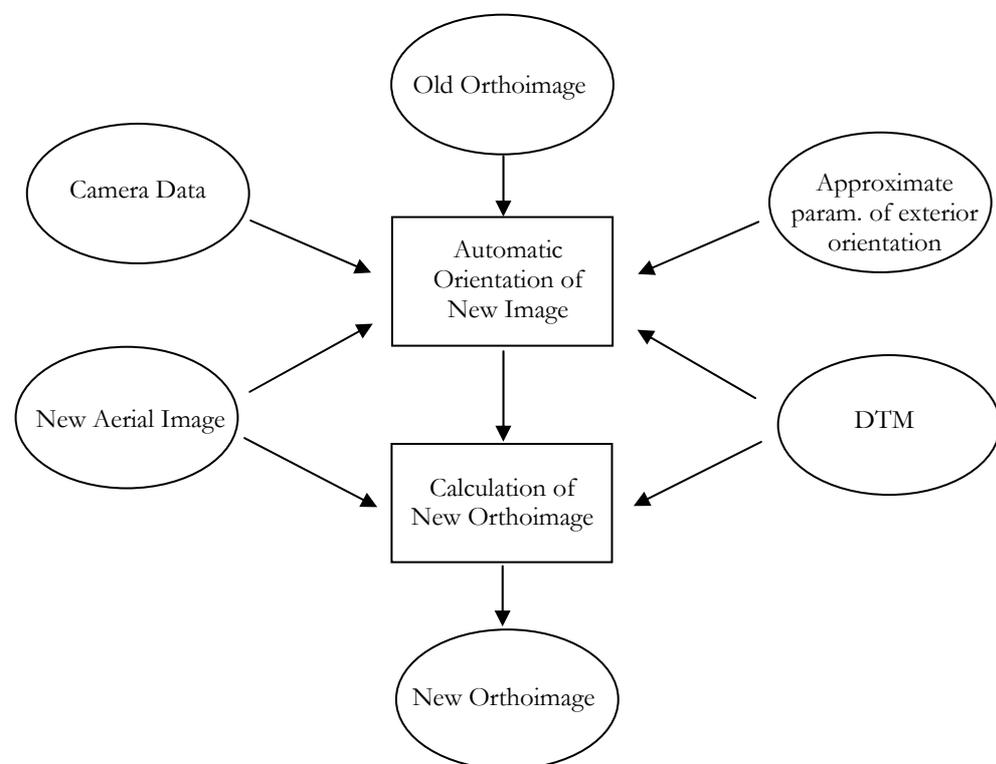


Fig. 2.1 Steps in the production of the next generation of orthoimages. Ellipses comprise input parameters, rectangles stand for calculation procedures (adapted from Höhle, 2003).

2.3.1 Goals of the investigation and the method overview

The starting point for an investigation of possible improvements of the chosen procedure of exterior image orientation was the detailed analyses of the method description and results published in (Höhle, 1999a, 1999b, 2001). The main goals for the investigation were summarised into following points:

- 1) Replacing a manual selection of control points both in the orthoimage and an aerial image by an automated procedure. A digital vector topographic map or database can be used for finding a position of suitable objects. The first experiments with this approach were published in (Höhle and Potůčková, 2001).
- 2) Checking the suitability of chosen patches for matching. The number of patches should be higher in comparison to the test described in the above mentioned literature.
- 3) Applying of hierarchical approach in order to decrease requirements for accuracy of approximation of orientation parameters.
- 4) Improving the accuracy of measurement in the images to subpixel range e.g. by means of least squares matching. The question is whether more accurate matching methods will have an effect on overall accuracy due to the time changes of chosen control points and their surroundings and the accuracy of the DTM.
- 5) Applying efficient methods for the detection of outliers based either on thresholding similarity measures or robust estimators described in chapters 1.2.1.5 and 1.3.
- 6) Checking of quality of the results by means of comparing of a newly derived orthoimage with an old orthoimage or other data set of sufficient accuracy, e.g. a topographic map. This step should be also carried out automatically.

Figure 2.2 shows the workflow of the suggested procedure. All steps were carried out by means of own developed programs except of creating an image pyramid, an orthoimage derivation, and measuring fiducial marks in an aerial image. Different parts of the software package ImageStation™ from Z/I Imaging were used for those purposes.

In the following chapter some characteristics of the used digital topographic maps and orthoimages are pointed out. A suitability of different objects for providing control information for orienting medium scale images (1:13 000 – 1:60 000) is also discussed.

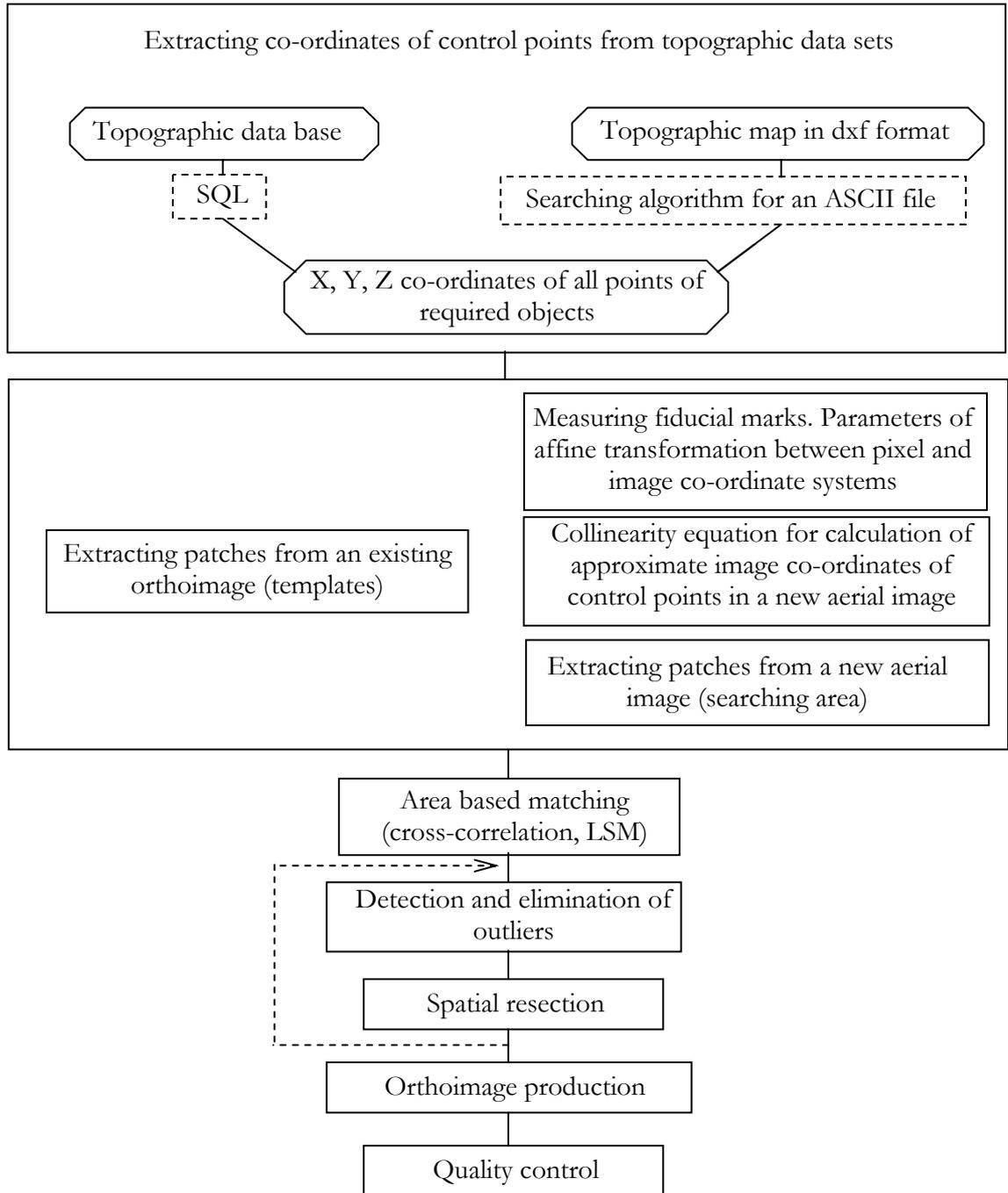


Fig. 2.2 Workflow of a tested method for automatic orientation of aerial images including orthoimage derivation and quality control. X, Y, Z co-ordinates of suitable objects are derived from a topographic database. Image patches are extracted from an existing orthophoto (templates) and a new aerial image (searching areas) and matched. Parameters of exterior orientation (EO) are calculated by means of spatial resection. It is carried out in an iterative process together with detection and elimination of outliers. A new orthoimage is derived and compared with the existing topographic orthoimage or map in order to evaluate its accuracy with regard to the existing data sets (adapted from Potůčková, 2003).

2.3.2 Combination of a topographic database and orthoimage for extracting control information

One of the basic requirements for topographic databases or geographic information systems is a good organisation of data. Each object has its code according to an object class and object type and is defined by a set of co-ordinates in a given reference system. Topology defines both the relations between geometric primitives that create an object (points and lines) and relations between objects, object classes and types. Depending on the database system, suitable commands are applied in order to extract information about the position of points or objects that can be used as control information for image orientation. The structured query language (SQL) represents an example of a language containing commands for a formulation of queries to a database [www.sql.org]. If the map is available only in a CAD system, it can be saved in the ASCII form and suitable objects can be found by means of a search program.

The question is what objects can be used as control information. As mentioned in (Höhle, 1999b) they should be suitable for area based matching and also be time invariant due to time interval between taking images. **Buildings** are an example of well defined and time invariant objects but their appearance differs due to a different viewing angle. Therefore they do not represent objects suitable for area based matching of large or medium scale images and a high geometric resolution (30 μ m and better). The elevations of building are not included in DTM. Thus, a 3D digital topographic map including building elevations or a DSM must be available in order to have correct heights for calculation of orientation parameters. On the other hand, if high accuracy is not required and a pixel size of an image is reduced to the level that a building is represented only by a few pixels, they can be easily used for matching purposes, e.g. for improving approximations of orientation parameters (hierarchical approach, see also chapter 2.4.2, point E).

Borders of **areal objects** as forests, fields, or lakes change a lot with time (in a level of meters) and therefore are not suitable for orienting medium or large scale images. Nevertheless, they can be useful for georeferencing of low-resolution satellite data. As mentioned in chapter 2.2.1 in large scale images **point like objects** as manhole covers and drain gratings can be also used. However, it is difficult to recognise such object in images of scale e.g. 1:15 000 and smaller (e.g. a manhole cover with a diameter of 0.5 m will cover about 2 x 2 pe^2 in an image 1:15 000 with a pixel size of 15 μ m).

Until now **road crossings** have been found as most suitable objects for the orientation of medium scale images. Roads have a relatively good contrast to the surrounding. Road crossings are usually flat areas and their accurate heights and positions are stored in the 3D topographic maps derived by stereophotogrammetric measurements. Using topologic relations, crossings surrounded by high vegetation or buildings that could cause problems in matching because of hidden areas, shadows, and time changes can be avoided. The amount of consolidated roads is usually high enough also in open land areas.

After the position of a suitable object, e.g. a centre of road crossing is extracted from a topographic database or map, an image patch consisting of $m \times n$ pel^2 where both m and n are odd numbers is cut from an orthoimage. It is the centre of the orthoimage patch which position has to be found in the aerial image by matching. In order to avoid shifts up to 0.5 pixel in each co-ordinate, not the position extracted from the map but the centre of the patch is considered as a reference point (see Fig. 2.3).

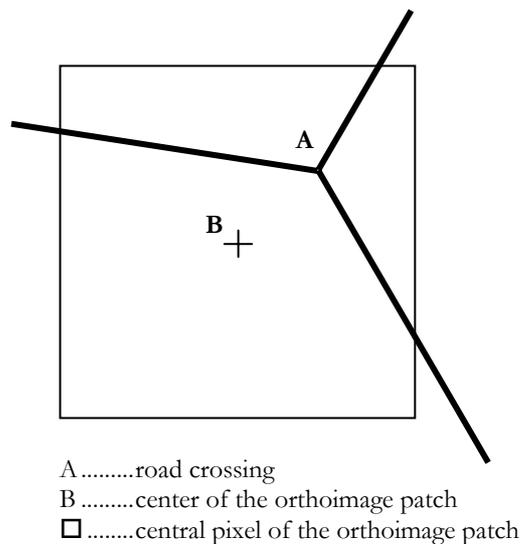


Fig. 2.3 Position of the road crossing A and a centre of the middle pixel of the orthoimage patch B. In order to avoid shifts in the position of the corresponding points in the orthoimage and the new aerial image, ground co-ordinates of the centre B should be used in calculation of spatial resection instead of co-ordinates of point A.

For calculating orientation parameters, an accuracy of control points that in this case corresponds to an accuracy of existing orthoimage must be known. In general it consists of two parts – an accuracy of orientation parameters of an aerial image and an accuracy of a DTM that were used for generating the orthoimage. According to (Honkavaara et al., 1999) an average standard deviation in the position of a single point in the orthoimage can be calculated

by means of formula (2.1). Assuming accurate control points, an accuracy of $\sigma_{x,y} = \sigma_0 = 0.3 - 0.4$ pel can be achieved by aerotriangulation with an automatic tie point measurement (Ackermann, 1996a). An accuracy of the DTM depends on the procedure used for its generation or derivation and is discussed in the chapter 3.

$$\sigma_{ortho} = \sqrt{\sigma_{ortho_DTM}^2 + \sigma_{ortho_ori}^2} = \sqrt{\left(\sqrt{\frac{2}{3}} \frac{a}{H} \sigma_{DTM}\right)^2 + \left(\frac{H}{c} \sigma_{ori}\right)^2} \quad (2.1)$$

σ_{ortho} standard deviation of a position of a point measured in an orthoimage

σ_{DTM} standard deviation of a height of a single point of a digital terrain model (DTM)

σ_{ori} a posteriori standard deviation of a position of a point measured in the image derived from aerotriangulation $\sigma_{ori} = \sqrt{2}\sigma_0$

σ_0 a posteriori standard deviation in image co-ordinates determined from aerotriangulation

c camera constant

H flying height

$2a$ image size

The height of the patch centre has to be interpolated from DTM. If a topographic map is available and a surrounding of a chosen point is flat, an area of 1 pixel can be considered as a horizontal plane with the height corresponding to the point from the map.

2.4 Test of the method

In this chapter, practical results with the method of an automated exterior orientation of an aerial image based on a 3D topographic map, an orthoimage, and DTM are reported. First, the data set used for all investigations is described. A discussion of all the steps as they were carried out and of the obtained results follows.

2.4.1 Data set description

The data used for this investigation are a part of a data set produced by the Danish Centre for Remote Sensing and the National Survey and Cadastre at the end of the year 1999. It covers area of approximately 112 km² in the surrounding of the town of Hvorslev (central Jutland, Denmark).

Vector data

The vector or map data have their origin in the Danish nation-wide three dimensional digital topographic map and the basic topographic database TOP10DK. It has been developed and maintained by the Danish National Survey and Cadastre [www.kms.dk]. According to the specifications it was created with the purpose of being:

- the basic database for the production of topographic maps
- the reference frame for other geographic registers
- the topographic base for geographic information systems

The level of detail corresponds to the topographic map 1: 10 000. Until now (June 2004), all the co-ordinates are calculated in the Universal Transverse Mercator (UTM) projection and referred to the European Datum 1950 (ED50). The updating of the database should be carried out in an interval of five years. All the objects in the database are specified as points, lines or areas. 51 object types are grouped into 8 classes. The database is topologically structured. The basic method for obtaining vector data is stereoplotting from images of the scale 1:20 000 – 1:25 000. The accuracy of well-defined points is 1 m in X (easting), Y (northing), and Z co-ordinates (TOP10DK specifications, 2001).

The available data set contains only a part of TOP10DK saved in the dxf (drawing interchange file) format that is an ASCII format developed for transfer of AutoCAD drawings to other CAD systems. Dxf files have a standard structure¹ that allows for a simple search for points with given attributes such as number of layer, colour, line style etc. All the TOP10DK objects were divided into layers according to object types. The layers got codes corresponding to the object type codes. Therefore, searching for points of a certain object type can be done by finding points in the layer named with the required object type code.

Imagery

Orthoimage

- Derived from a B&W aerial photo 1:25 000, scanned with the pixel size of 25 μm
- Date of photography: 8th April 1995
- Ground sampling distance 0.625 m
- Specification of DTM see below
- Average accuracy $\sigma_{\text{O_old}}=1.00 \text{ m} \approx 1.6 \text{ pel}$ (calculated according to formula 2.1 for $a=11.5 \text{ cm}$, $c=15 \text{ cm}$, $H=3750 \text{ m}$, $\sigma_0=10\mu\text{m}$ and $\sigma_{\text{DTM}}=1.5 \text{ m}$)

¹ http://www.autodesk.com/techpubs/autocad/acad2000/dxf/dxf_format.htm

Aerial image

- B&W photo, scale 1:25 000, scanned with the pixel size of 25 μm
- Date of photography: 9th April 1999

Both of images were taken with a wide-angle aerial camera ($c=15$ cm). Each of images covers about 33 km^2 and they overlap of about 88 %. They are saved in tif format.

Digital terrain model

DTM covering the test area is represented by a square mesh regular grid of 25 m. It is a section of a national DTM that was established in Denmark in late 1980's and was generated from digitised contour lines with 5 m interval. Later on the model was improved by means of contour lines with the interval of 2.5 meters and heights from TOP10DK. The point accuracy of the model is quoted to 1 or 1.5 meters (Larsen, 1998).

The differences between DTM and TOP10DK within the test area were checked by a comparison of heights of 60 road crossings. The obtained root mean squared error $\text{RMSE}_{dZ}=1.2$ m is in a good correspondence with an expected standard deviation $\sigma_{dZ}=(\sigma_{\text{TOP10DK}}^2+\sigma_{\text{DTM}}^2)^{1/2}=(1^2+1.5^2)^{1/2}\text{m}=1.8\text{m}$. A systematic shift $dZ_{\text{TOP10DK-DTM}}=0.65$ m was discovered. It corresponds to a shift in position of 0.65 m at the corner of an image 1:25 000 taken with a wide-angle camera ($c=15$ cm). In the applied method the DTM is not used for the derivation of orientation parameters, but only for the derivation of a new orthoimage. Therefore the shift should appear after comparison of the new orthoimage and the topographic map.

2.4.2 Steps in the application of the method and results

A) Extraction of control information from the digital vector map

TOP10DK objects were evaluated with respect to their suitability as control information for an orientation of an image at the scale of 1:25 000. The definition, appearance in the image, size, surroundings, and distribution over the test area were the main criteria. Finally, crossings of three road types were chosen – roads over 6 m wide (code 2115), road 3-6 m wide (code 2122), and other roads (code 2123). The road network is shown in Fig. 2.4a.

The road crossing extraction from the dxf file was done by own developed program ***crossings.pas*** (developed in TurboPascal v. 7.0). The idea of the search is based on

TOP10DK structure. Each of the lines is registered by its start and end point. The road crossing is a point where at least three lines start or end. Therefore at least three points with the same co-ordinates appear in the dxf file. If all of them have one of selected codes (2115, 2122, or 2123) the co-ordinates are saved into the output file. After extracting of all road crossings their number was thinned out. The test area was overlaid with a grid of 500 m x 500 m and a criterion of one point per a grid square was applied. Only points in the area of the orthoimage and aerial image were taken into account. All developed programs for extracting road crossings and thinning out the data can be found in Appendix C. In order to avoid road crossings surrounded by high vegetation, the points were superimposed on the orthoimage and those in forest areas taken from the list manually. This step could be easily automated if the full database was available (a query whether a point lies inside or close to forest polygons). Programming of such a step based on information from dxf file would be time consuming and not completely relevant to the topic of the thesis and therefore a manual solution was applied. Finally, 57 control points were used for further calculations (Fig.2.4b).

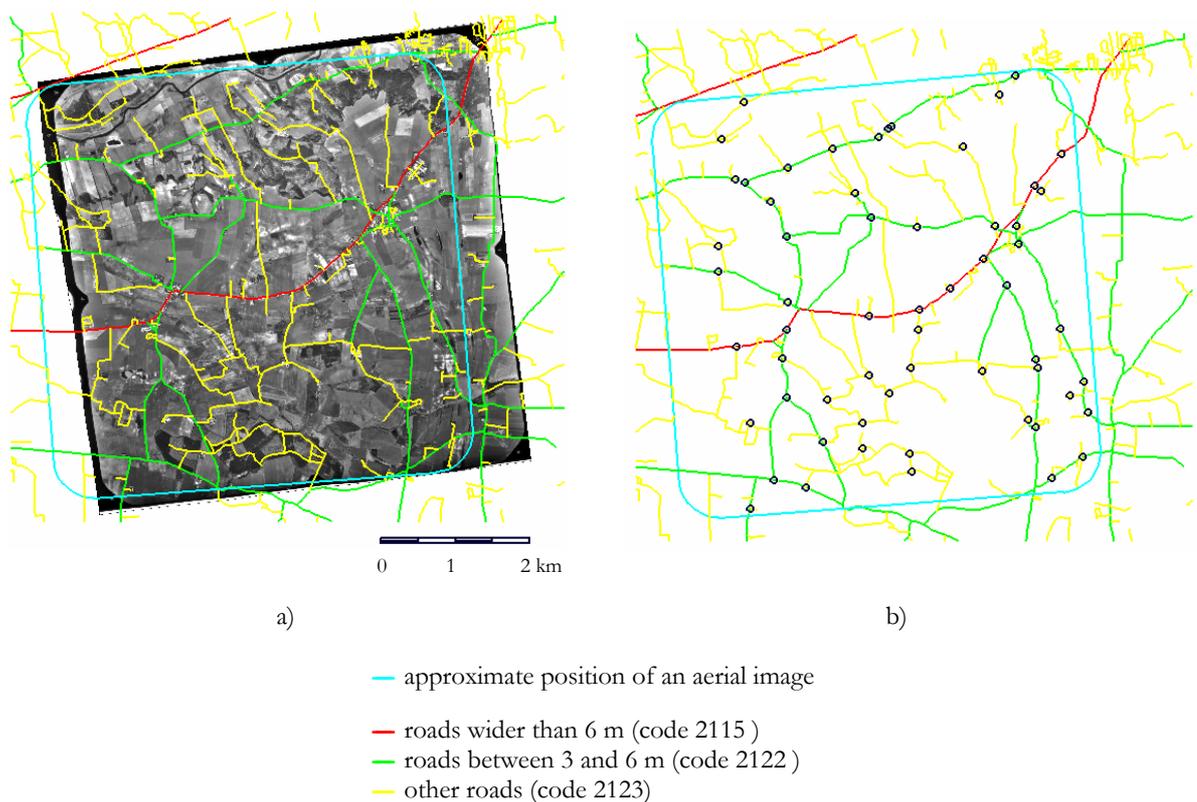


Fig. 2.4 a) Road network in the test area superimposed on the existing orthoimage. b) Road network with 57 selected control points. The approximate position of an aerial image is marked with a light blue line.

B) Verifying of the relative accuracy between the orthoimage and the topographic map

Errors in the orientation parameters in the images used for a derivation of the existing orthoimage and topographic map as well as inaccuracy of measurement in stereomodels or errors in DTM can cause both local and overall shifts between the orthoimage and the topographic map. The derived orthoimage patches will then be moved from the centre of road crossings which can cause troubles with correlation due to high objects in neighbourhood, low contrast in field areas or insufficient structure, e.g. correlation along a line object (the template covers only a road, not a crossing).

The fit between the existing orthoimage and the vector map was checked visually. A scaling problem was visible at the edges of the orthoimage. It could be caused by a systematic shift between TOP10DK and DTM heights. Therefore 25 well-distributed points were measured. The mean differences $M_x = 0.05$ m and $M_y = 0.03$ m do not show any systematic shift in position. Root mean square errors calculated from differences $RMSE_x = 1.01$ m and $RMSE_y = 1.11$ m correspond to accuracy of manual measurement of road crosses (1-2 pel). However, scale problems are visible from error vectors in Fig. 2.5a.

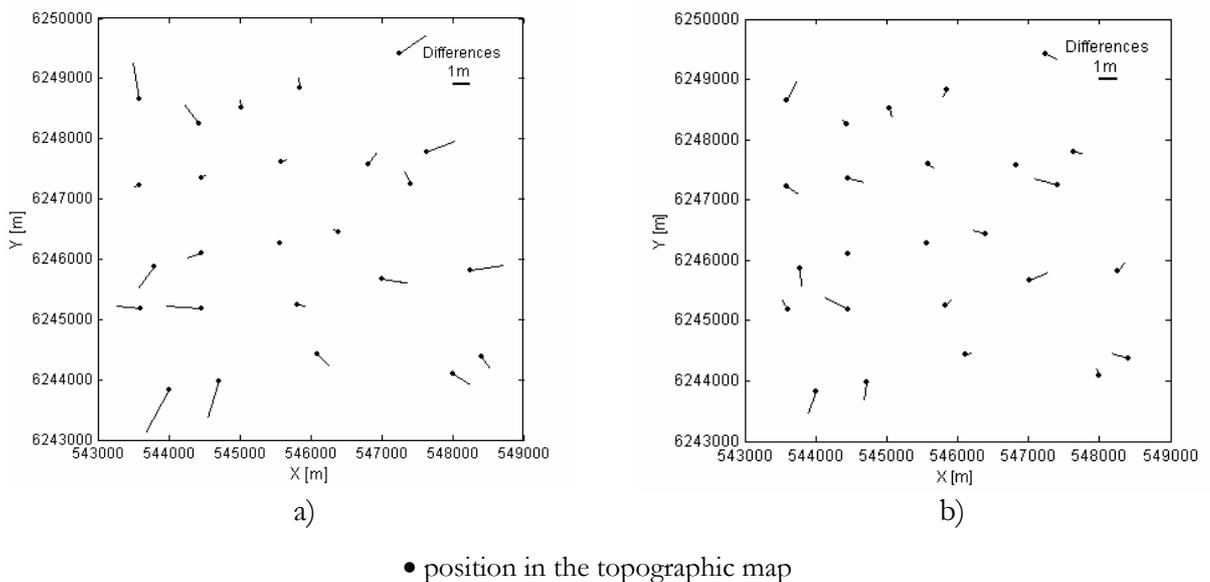


Fig. 2.5 a) Differences between TOP10DK and the existing orthoimage measured in 25 road crosses. From the size and directions of vectors scale difference between two data sets is visible. b) Differences after a conform transformation.

A conform transformation from the orthoimage to the map revealed a small rotation of 1 mgon and a standard deviation $\sigma_{x,y} = 0.68$ m (improvement of about 30% in comparison to

$RMSE_{x,y} \approx 1$ m presented above). The scale factor $k=0.999476$ (52.4 cm/km) gives the explanation to relatively large errors vectors at the corners of the orthoimage.

For derivation of orientation parameters the position of the control point is taken from an orthoimage and the height from TOP10DK. The orientation parameters of the new aerial image will therefore be influenced by the scale of existing orthoimage (see also chapter 2.5). The absolute accuracy of a new orthoimage derived from these orientation parameters cannot be improved with respect to the existing one. However, it is the fit between the new and the old orthoimage or the map that will be evaluated and that is also important when e.g. new changes in the landscape must be included into existing data sets.

C) Extraction of orthoimage patches

The co-ordinates of an upper left corner and a ground sample distance (gsd) of the existing orthoimage are given in tfw (tif word file) format. The relation between a reference system and pixel co-ordinate system of the orthoimage is shown in Fig. 2.6.

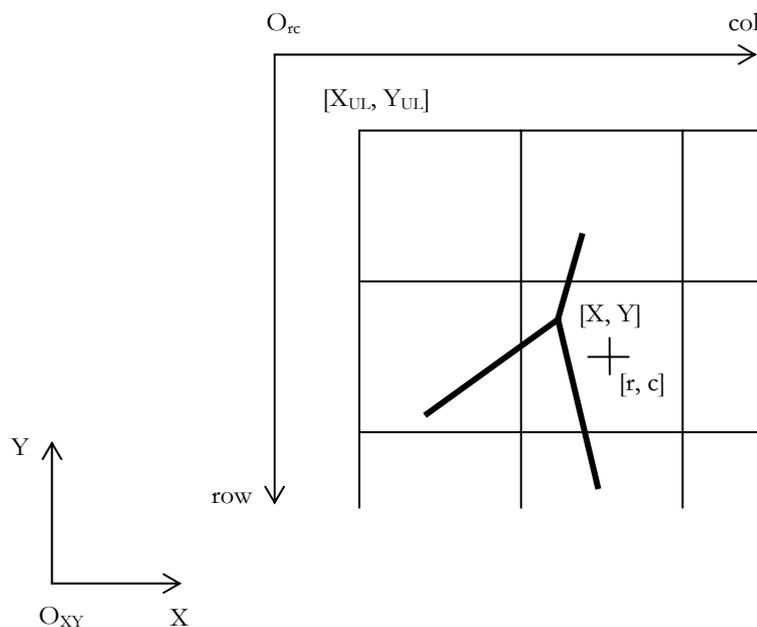


Fig. 2.6 Relation between the reference co-ordinate system $[O_{xy}, X, Y]$ and orthoimage pixel co-ordinate system $[O_{rc}, row, col]$. X_{UL}, Y_{UL} are co-ordinates of the upper left corner of the orthoimage in the reference system. The origo of the pixel system is shifted $0.5gsd$ out from the image.

The pixel co-ordinates of the centre of the image patch containing a road crossing with co-ordinates X and Y are calculated according to formula 2.2.

$$r = \text{round}((Y_{UL} - Y) / \text{gsd} + 0.5) \tag{2.2}$$

$$c = \text{round}((X - X_{UL}) / \text{gsd} + 0.5)$$

X_{UL}, Y_{UL} co-ordinates of the upper left corner of the orthoimage in the reference co-ordinate system

gsd ground sample distance

X, Y co-ordinates of the road crossing in the reference co-ordinate system

r, c co-ordinates of the centre of the pixel containing road crossing in the pixel co-ordinate system

round function rounding a number to the closest integer value

The width of roads in the test area is in the interval from 2 m to 9 m. The measurements in the orthoimage showed that in single cases the road crossing width goes up to 15 m which corresponds to 24 pixels. Road crossings themselves do not have enough texture for automatic measurement by means of correlation. Their close surrounding has to be included. As a starting point, the size of image patches, templates, was chosen $31 \times 31 \text{pel}^2$ which corresponds to area of $19.4 \times 19.4 \text{m}^2$. Examples of templates are shown in Fig. 2.7.

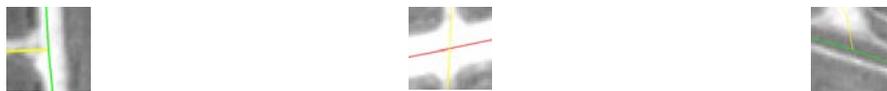


Fig. 2.7 Examples of orthoimage patches (templates) $31 \times 31 \text{pel}^2$ with superimposed line intersection from the topographic map.

D) Interior orientation of the aerial image

The camera calibration report was a part of the data set. Values of the camera constant, the image co-ordinates of the principal point and radiometric lens distortions were used as input parameters of the developed resection program (Appendix A). For finding transformation parameters from the pixel to image co-ordinate system and vice versa, pixel co-ordinates of fiducial marks were measured manually in the software package ImageStation™ Digital Mensuration from Z/I Imaging. Image co-ordinates of fiducial marks were available from the camera calibration report. Parameters of affine transformation were calculated in the MATLAB® script *aff_trf.m* (see Appendix C). Standard deviation of $\sigma_{x,y} = 7 \mu\text{m}$ was achieved. The measurement of fiducial marks and the computation of parameters of the affine transformation were done for three levels of derived image pyramid (see point E). The

scanner calibration report was not available. It is assumed that most of the errors were minimised by the affine transformation. The remaining irregular errors (at the level of 2-3 μm) do not have so high importance for images with a geometric resolution of 25 μm .

E) Approximations of exterior orientation parameters of the aerial image

The approximations of orientation parameters are needed in order to find centres of search image patches in an aerial image. The knowledge about their accuracy is important for defining the size of the search image patches. The approximations must be as accurate that the ‘control point’ displayed in the template derived from the orthoimage is contained in the search patch. The following example shows what shifts in the image are caused by an error of 10 m in co-ordinates of the perspective centre and 0.2 gon in rotations:

Flying height $h=3750$ m

Camera constant $c=150$ mm

Pixel size 25 μm

Error in the orientation parameters

Shift in the image due to orientation parameters
(calculated for $x'=y'=115$ mm)

$e_x = e_y = 10$ m

$dr'_x = dr'_y = 16$ pel

$e_z = 10$ m

$dr'_z = 12$ pel

$e_\omega = e_\phi = 0.2$ gon

$dr'_\omega = dr'_\phi = 32$ pel

$e_\kappa = 0.2$ gon

$dr'_\kappa = 20$ pel

It is not likely that all the errors would appear with the same size. Nevertheless, if shifts of e.g. 30 pixels are considered and the size of the template is 31 x 31 pel², the size of the search area must be at least 31+2*30=91 pixels. If correlation is used, both the possibility of mismatches and the calculation time increases with an increasing search area. In a given example, $(91-31+1)^2=3721$ calculations of correlation coefficient per one control point is necessary to find a position of the best fit. As mentioned in the chapter 2.2.1, a hierarchical approach gives a possibility for improvements of the approximations.

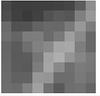
There was no information about orientation parameters of the new aerial image in the given data set. To get some starting values, nine points were measured manually and orientation parameters were obtained by means of spatial resection (see Tab. 2.1).

X_0 [m]	Y_0 [m]	Z_0 [m]	ω [gon]	φ [gon]	κ [gon]
545 600	6 246 542	3809	0.011	-0.603	5.200

Tab. 2.1 Orientation parameters of a new aerial image obtained by means of spatial resection. Nine control points (centres of road crossings) were measured manually.

In order to simulate a real situation when the approximations of orientation parameters are known from GPS measurement or flight planning, the calculated orientation parameters were changed as one can see in Tab. 2.3. The co-ordinates X_0 and Y_0 were shifted of about $\Delta X = \Delta Y = 50$ m. Z_0 was determined from a camera constant, approximate image scale and an average height of the terrain that was about 50 m, e.g. $\Delta Z = 9$ m. Under assumption of an image with the vertical optical axis, the tilts ω and φ were set to zero (e.g. $\Delta\omega = 0.02^\circ$, $\Delta\varphi = 0.6^\circ$) and κ angle changed of about $\Delta\kappa = 2^\circ$.

An image pyramid was created for both the orthoimage and the aerial image by means of the software package ImageStation™ Raster Utility – Overview Utility from Z/I Imaging. The sampling increment factor of 2 was used. The Gaussian filter was applied for each level of the pyramid before sampling. A determination of parameters of exterior orientation of an aerial image for levels 2 – 4 was carried out. For each pyramid level co-ordinates of 8 – 12 well-distributed control points were manually found in the TOP10DK. For higher levels of the image pyramid (4 and 3) only crosses of roads at least 6 m wide and houses were used. With this respect, it was advantageous that the roofs of houses were registered in TOP10DK and therefore all control points were introduced with correct heights. Templates and search areas were extracted from orthoimage and aerial image pyramid levels (see also points C and F). Tab. 2.2 shows examples of templates derived from the orthoimage.

Pyramid level	Pixel size [μm]	Ground sample distance [m]	Template size [pel ²]	Examples of control patches (Object pyramid)	
				house	road crossing
4	400	10.0	9 x 9		
3	200	5.0	11 x 11		
2	100	2.5	17 x 17	-	

Tab. 2.2 Example of image and object pyramid used for improving approximations of orientation parameters of the aerial image.

The position of the best fit between image patches was found by means of cross correlation. The orientation parameters for all three pyramid levels were calculated using spatial resection. Due to small amount of control points the blunders were excluded by data snooping (see chapter 1.3.1). If the importance of an excluded point was high (with respect to point distribution), a new control point was found and the calculation was repeated. The results of orientation of three image pyramid levels are summarised in Tab. 2.3.

Pyramid level	Pixel size [μm]	Orientation Parameters					
		X ₀ [m]	Y ₀ [m]	Z ₀ [m]	ω [gon]	φ [gon]	κ [gon]
start values		545 650	6 246 500	3800	0	0	7
4	400	545 609	6 246 543	3812	-0.059	-0.576	5.067
3	200	545 599	6 246 543	3810	-0.021	-0.660	5.188
2	100	545 601	6 246 540	3812	0.020	-0.601	5.200

Tab. 2.3 Approximation of orientation parameters obtained for three levels of the image pyramid.

The orientation parameters obtained from levels 3 and 2 did not differ more than 3 m in coordinates of the projective centre and 0.06^s in rotations. Taking each of these differences separately, they correspond to shifts up to 8 pel in the level with original pixel size of 25μm. Moreover, the obtained approximations from level 2 fit also to the result based on manual measurement (Tab. 2.1). Therefore calculation of exterior orientation parameters of the first level of the image pyramid was not necessary and approximations from the second level were used directly for orientation of the level with the highest pixel size.

F) Extraction of search patches from the aerial image

Image co-ordinates of all control points derived were calculated by means of collinearity equations (formula 1.3). Small values of lens distortions and position of the principal point were neglected. The values derived for the second image pyramid level (Tab. 2.3) were used as approximations of orientation parameters. The image co-ordinates were transformed into a pixel co-ordinate system using the inverse affine transformation which parameters were derived in step D. Pixel co-ordinates were rounded to closest integer value, i.e. they correspond to the centre of the pixel which also became the centre of the search area.

The accuracy of approximation of orientation parameters can be assumed better than 5 m (≈0.13%h) in position and 0.1^s in rotations. A single error of the size of 5 m or 0.1^s can cause a shift of 8 pel at the edge of the image. Due to cumulating of errors, a possibility of a shift of 15 pel was assumed. Therefore, the size of search area was chosen as 31+2*15 pel = 61 pel.

G) Matching image patches

Based on investigations presented in the chapter 1.2.1.5, the correlation coefficient was chosen as a similarity measure. In connection with the investigated method, an application of matching techniques allowing for subpixel accuracy have been mentioned in literature (Höhle, 2003) but it has not been applied so far, except of (Potůčková, 2003). Therefore the matching was running in two steps. First, the position of the best fit was calculated by means of the correlation coefficient and the found position was used as an approximation for least squares matching (LSM). The radiometric corrections were carried out prior to LSM by means of the linear function. Geometric models with 2, 4, and 6 parameters were applied. The best results were expected from the 4 and 6 parameter models due to rotations and small scale differences between the orthoimage and the aerial image. In case of failure of the LSM, a value obtained by means of correlation coefficient was used for the calculation of resection. In order to get a full overview of results from different subpixel methods, a simple approach based on an approximation of correlation surface with polynomial of the second order was also applied (see Appendix B.1). The calculation was carried out by means of a MATLAB[®] script ***ABM.m***. A text file with the calculated positions of the best fit in the pixel co-ordinate system and relevant attributes as correlation coefficient, standard deviations of shift parameters of LSM, etc. was an output of the program. The reports from the calculation and the program description can be found in Appendix C.

H) Calculation of exterior orientation parameters

The pixel co-ordinates obtained from image matching were transformed into image co-ordinate system. First, the image co-ordinates obtained by means of cross-correlation were used for calculation of orientation parameters. Because the correlation coefficient for all the points was higher than 0.5, none of points was excluded as a mismatch prior to adjustment. The calculation of the orientation parameters was done by means of spatial resection with robust adjustment (see Appendix A). A priori standard deviations were set to $\sigma_0=12.5 \mu\text{m}$ ($\approx 0.5 \text{ pel}$), $\sigma_x=\sigma_y=\sigma_{\text{O_old}}/\sqrt{2}=0.7 \text{ m}$, $\sigma_z=1 \text{ m}$. Eight discovered outliers were checked in the orthoimage. It was found out that in some of them the size of the template was not big enough. Therefore all points were recalculated with a template size of $41 \times 41 \text{ pel}^2$. The co-ordinates of six points changed more than 1 pixel. The correlation coefficient of one point reached the value of 0.25 only and therefore this point was excluded from further calculations.

The calculation of resection using the Danish robust adjustment method was carried out for all similarity measures with new co-ordinates of the seven points mentioned above. A priori

standard deviations were set again to $\sigma_x=\sigma_y= 0.7$ m, $\sigma_z=1$ m. The standard deviation of the measurements in the image $\sigma_0=12.5$ μm (≈ 0.5 pel), was changed to $\sigma_0=8$ μm (≈ 0.3 pel) for the methods with the calculation in the subpixel range. The standard deviations were set equal for all points. In order to get reliable results the control points should be evenly distributed over the whole image. Therefore the centre of gravity of all the control points that were not weighted down in the robust adjustment was calculated. An area formed by those control points was compared with an area of the entire image (23×23 cm^2).

I) Results

The orientation parameters calculated for different matching methods are summarised in Tab. 2.4. The resection calculation reports can be found in Appendix C.

Matching method	Orientation Parameters					
	X_0 [m]	Y_0 [m]	Z_0 [m]	ω [gon]	ϕ [gon]	κ [gon]
r	545 599.53	6 246 544.19	3810.92	-0.018	-0.607	5.207
r_sub	545 599.60	6 246 544.19	3810.92	-0.018	-0.606	5.207
LSM_2	545 599.76	6 246 544.32	3810.87	-0.019	-0.604	5.208
LSM_4	545 600.39	6 246 543.75	3811.00	-0.011	-0.594	5.211
LSM_6	545 599.33	6 246 543.36	3810.77	-0.006	-0.616	5.212

r cross correlation
r_sub cross correlation with subpixel accuracy
LSM_x least squares matching with x geometric parameters

Tab. 2.4 Orientation parameters obtained for five matching methods.

The maximal differences in the orientation parameters up to 1 m in the co-ordinates X_0 and Y_0 and 22 mgon in the ϕ angle are present. The differences in the values Z_0 do not exceed 20 cm and the rotations ω and κ do not differ more than 5 mgon. The variances are caused by different positions of control points found in the aerial image by means of different matching techniques and also by number of outliers excluded in the robust adjustment. Standard deviations of orientation parameters achieved by different matching methods have almost the same magnitude (differences between standard deviations are smaller than 12 cm in the position and 1.5 mgon in the rotations). Tab. 2.5 gives an overview of standard deviations of calculated parameters and a posteriori standard deviation of measurement in the image σ_0 .

Matching method	Accuracy of orientation parameters						
	σ_{x_0} [m]	σ_{y_0} [m]	σ_{z_0} [m]	σ_{ω} [mgon]	σ_{ϕ} [mgon]	σ_{κ} [mgon]	σ_0 [μm]
r	0.57	0.55	0.21	8.2	8.2	2.6	11
r_sub	0.56	0.53	0.21	8.0	8.3	2.5	7
LSM_2	0.57	0.55	0.21	8.2	8.5	2.5	7
LSM_4	0.52	0.49	0.19	7.4	7.7	2.3	6
LSM_6	0.64	0.60	0.24	8.9	9.6	2.8	8
Mean	0.57	0.54	0.21	8.1	8.5	2.5	-

Tab. 2.5 Standard deviations of orientation parameters and a posteriori standard deviation of measurement in the image σ_0 obtained for five matching methods.

The mean values of standard deviations σ_{x_0} , σ_{y_0} and σ_{ω} , σ_{ϕ} correspond approximately to 23 μm (≈ 0.9 pel) respectively 20 μm (≈ 0.8 pel) in the image. It is approximately two times worse than accuracy achievable by aerotriangulation, e.g. 0.4 pel (Ackermann, 1996a). The mean standard deviation in height is small (0.005%h). Also the standard deviation σ_{κ} does not cause shifts bigger than 4 μm (0.2 pel) at the corners of the image.

The lowest standard deviations of the orientation parameters were achieved for least squares matching method with four geometric parameters. Comparing the results of matching and resection, most of points with high residuals can be characterised by a high number of iterations or a failure in LSM. Such points could be excluded prior to calculation of resection or their a priori weights could be set lower (e.g. 0.5 pel for points where LSM failed and results of cross correlation were used).

As Tab. 2.4 shows, the orientation parameters - obtained by means of the rotation invariant matching methods (LSM4 and LSM6) - differ from the parameters obtained by application of methods dependent on rotation (r, r_sub, and LSM2). The subpixel measurement based on the rotation dependent methods did not bring any improvement in accuracy of obtained parameters with respect to applying cross-correlation only. For further calculations orientation parameters obtained for measurements of control points by cross-correlation and LSM with four geometric parameters (LSM4) are used as representatives of both groups.

The check of obtained orientation parameters could be done by comparison with orientation data calculated by a more accurate method, e.g. aerotriangulation or by using well-defined check points. As mentioned above, no information about the orientation was available. The

check points from topographic map could be chosen but their measurement in the aerial image is rather inaccurate (1-2 pixels). As mentioned at the beginning of the chapter 2.3, the investigation was done with the purpose to find out whether the orientation method is sufficient for orthoimage production. If the new orthoimage is derived and the accuracy of the DTM is known, it is possible to judge about the errors in the orientation parameters from the comparison of the new orthoimage with existing data. At the same time it will be possible to find out whether applying subpixel methods brings any improvements with respect to accuracy.

2.5 Quality control of the new orthoimage

After derivation of orientation parameters, a new orthoimage can be generated. For map updating and change detection purposes, a relative accuracy and fit to existing data is important. Thus, quality control of the new orthoimage must be carried out. Check points in the new and old orthoimages as well as in the topographic map are measured for this purpose. In case of image data check points can be replaced by check image patches. It avoids less accurate manual measurement. In case of map data, vector to raster conversion or object extraction from an orthoimage must be carried out in order to automate the process (an overview with references to literature can be found in (Heipke, 2004)). Developing and application of such procedures are beyond the concern of this thesis. Comparison with the map data will therefore be carried out only manually.

The new orthoimages were derived in the ImageStation™ Base Rectifier program. The ground sampling distance was set equal to the value used in the existing orthoimage, here 0.625 m. In order to minimise errors from height interpolation, the minimal possible distance between anchor points of 2 pixels and the bilinear interpolation method were chosen. Depending on the method used for matching image patches for finding position of control points in the aerial image, orthoimages are called O_LSM4 (LSM with four parameters) and O_r (correlation coefficient).

In the next chapter the procedure and results of the comparison of the orthoimages O_LSM4 and O_r with the existing orthoimage are described. A comparison of the orthoimage O_LSM4 with the topographic map follows. An explanation of discovered errors caused by differences in heights between TOP10DK and the DTM is also included.

2.5.1 Automated comparison with the existing orthoimage

Automated comparison between two orthoimages can be carried out in three steps:

A) *Extracting orthoimage patches* (see also point C, chapter 2.4.2).

122 road crossings were used as check points. Their co-ordinates were extracted automatically from the topographic map. The points in forest areas were excluded manually. Templates of 31 x 31 pel² were derived from an existing orthoimage and search areas of 51 x 51 pel² were extracted from a new orthoimage. It allows for shifts up to 10 pel. Such shifts should not appear in orthoimages with an average accuracy about 1 m (1.6 pel).

B) *Matching orthoimage patches*

Matching was carried out by means of cross correlation and LSM with four geometric parameters. Radiometric adjustment was carried out prior to matching. The geometric model with two parameters (shifts) should be sufficient. Four parameters were chosen in order to handle possible rotations between orthoimages due to different orientation parameters.

C) *Evaluation of the results*

Co-ordinates of calculated positions of the best fit in the new orthoimage were compared with co-ordinates of the centres of the templates from the existing orthoimage. The existing (old) orthoimage was considered as the reference data set. By 'error' a difference 'new_ortho – old_ortho' is meant in this chapter.

According to the normal distribution, standard deviations $\sigma_{O_X} = \sigma_{O_Y} = 1.09$ m (≈ 1.7 pel) and differences smaller than $3\sigma_{O_X} = 3\sigma_{O_Y} = 3.27$ m (≈ 5.2 pel) in both X and Y co-ordinates were expected. The values of σ_{O_X} and σ_{O_Y} were derived from accuracy of orthoimages using formula 2.3.

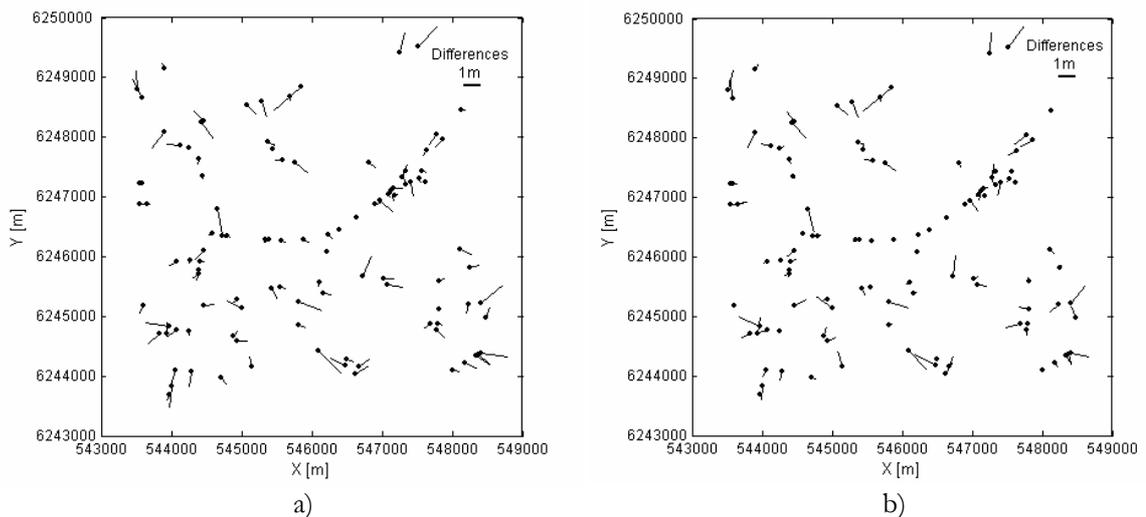
$$\sigma_{O_X}^2 = \sigma_{O_Y}^2 = (\sigma_{O_{Old}}^2 + \sigma_{O_{New}}^2)/2 \quad (2.3)$$

$\sigma_{O_X}, \sigma_{O_Y}$ standard deviations in differences between co-ordinates of check points measured in the new and existing orthoimages by means of area based matching

$\sigma_{O_{Old}}, \sigma_{O_{New}}$ average positional accuracy of the existing and new orthoimage

Using formula 2.1 an average standard deviation of the new derived orthoimage was calculated as $\sigma_{O_New} = 1.18$ m. A standard deviation of 0.5 m that corresponds to RMSE values in coordinates X and Y obtained for control points in spatial resection was taken as an estimation of accuracy coming from orientation. According to the chapter 2.4.1 $\sigma_{O_Old} = 1.00$ m.

First, orthoimage O_LSM4 was evaluated. Mean differences $M_X=0.33$ m (≈ 0.5 pel), $M_Y=-0.49$ m (≈ 0.8 pel) and root mean square errors $RMSE_X=1.38$ m (≈ 2.2 pel) and $RMSE_Y=1.96$ m (≈ 3.1 pel) were obtained. After excluding 16 points with differences $dX > 3RMSE_X$ and $dY > 3RMSE_Y$, remaining 106 points revealed root mean square errors $RMSE_X=0.55$ m and $RMSE_Y=0.64$ m. It corresponds approximately to 1 pel in both coordinates which is about 70% better than expected accuracy ($\sigma_{O_X} = \sigma_{O_Y} = 1.7$ pel). Mean differences showed small systematic shifts $M_X=0.13$ m (≈ 0.2 pel) and $M_Y=-0.15$ (≈ 0.2 pel). Fig. 2.8a shows the magnitude and direction of differences.



- position in the existing orthoimage

Fig. 2.8 a) Differences in position of 106 road crossings in the existing orthoimage (reference) and the new orthoimage O_LSM4. b) Residuals after conform transformation from the new to the existing orthoimage.

Vectors in Fig. 2.8a show that the scale of two orthoimages differs. Applying a projective transformation changed neither the size of the vectors ($RMSE_X=0.48$ m and $RMSE_Y=0.62$ m) nor their direction. After applying conform transformation the same root mean square errors were obtained ($RMSE_X=0.49$ m and $RMSE_Y=0.62$ m). The scale factor $k=0.999903$ (9.7 cm/km) explains a small decrease of the error vectors, see Fig. 2.8b. The rotation parameter was small, only 3 mgon (4 μ m in the corner of the image). It proves that there are not any

systematic errors in the orientation parameters of the new image with respect to the old one. The found differences must be therefore a result of matching or individual height differences between the DTM that was used for derivation of orthoimages and TOP10DK from which the heights of control points were taken.

Based on image matching reports in appendix C, the accuracy of 0.4 pel (≈ 0.25 m) or better in both co-ordinates can be expected when matching orthoimages of different date using LSM. Therefore attention was paid to changes in orientation parameters in relation to heights of control points. A new calculation was carried out. Heights of control points were interpolated from DTM instead of using heights from TOP10DK. The resection report can be found in Appendix C. A new orthoimage was derived and compared with the existing one in the same way as before. Root mean square errors $RMSE_X=0.29$ m and $RMSE_Y=0.32$ m corresponding to 0.5 pel and mean values $M_X=-0.05$ m and $M_Y=-0.09$ corresponding only to 0.1 pel were achieved (calculated from 115 points). It proves that the presence of height differences between the DTM and topographic map is the reason for differences in the position of corresponding points in two orthoimages. The problem is discussed in detail in the chapter 2.5.2. The size and direction of differences is shown in Fig. 2.9.

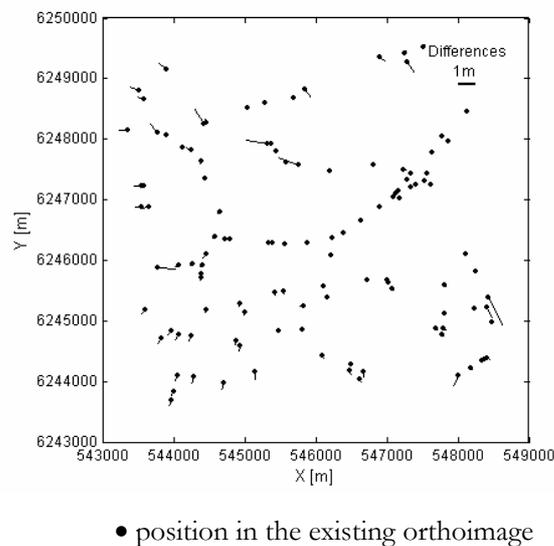
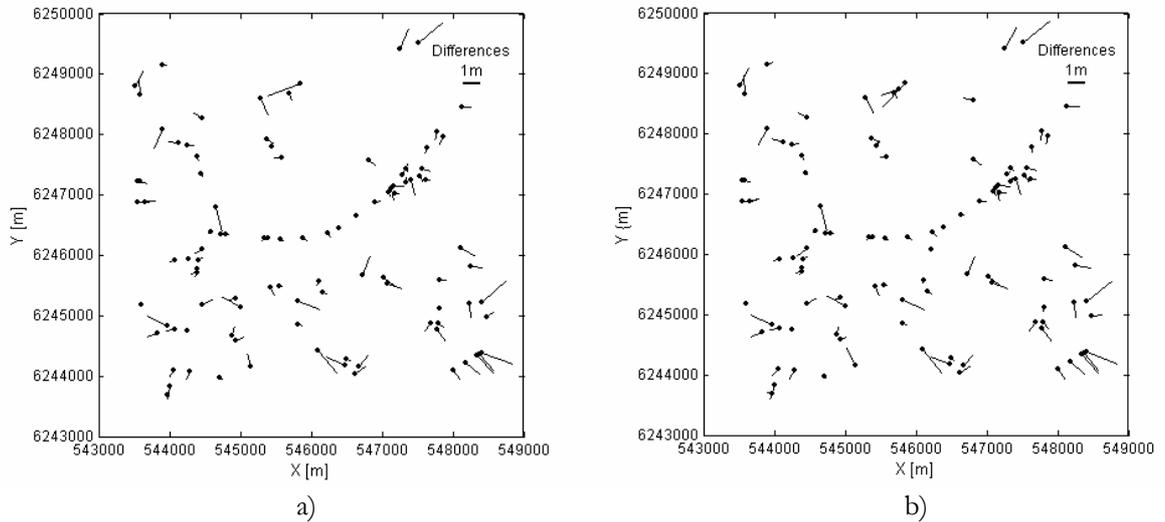


Fig. 2.9 Differences in position of 115 road crossings in the existing orthoimage (reference) and the new orthoimage O_LSM4 when DTM height were used in the calculation of orientation parameters of the new image.

Next, a comparison between the orthoimage O_r and the existing orthoimage was carried out. It revealed root mean square errors $RMSE_X=0.59$ m and $RMSE_Y=0.66$ m corresponding to 1 pel and systematic shifts $M_X=0.17$ m (≈ 0.3 pel) and $M_Y=-0.12$ (≈ 0.2 pel). New orientation

parameters using DTM heights were also calculated and a new orthoimage derived. But in this case, a better fit with the existing image was not achieved ($RMSE_x=0.60$ m, $RMSE_y=0.65$ m, $M_x=0.16$ m and $M_y=-0.15$ m). Fig. 2.10 shows graphically that the due to lower accuracy of matching the size and direction of vectors does not differ when heights from the topographic map and DTM were used for derivation of orientation parameters.



- position in the existing orthoimage

Fig. 2.10 Differences between the new orthoimage O_r and the existing one when heights from a) topographic map and b) DTM were used for calculating orientation parameters of the new aerial image.

2.5.2 Comparison with the topographic map

The results in the chapter 2.5.1 showed that when heights from the topographic map were used for derivation of orientation parameters, RMSE obtained for orthoimages O_{LSM4} and O_r did not differ more than 0.1 pel. Such a value can be neglected if checkpoints are measured manually which is the case of investigations carried out in this chapter. Therefore only orthoimage O_{LSM4} is used for all the calculations. The topographic map is considered as the reference data set. By ‘error’ a difference ‘map - orthoimage’ is meant in this chapter.

The comparison between the topographic map and a newly produced orthoimage was done by manual measurement of 25 well-distributed points. The same points that were used for the comparison of the topographic map and the existing orthoimage were chosen. The obtained differences should be within an expected accuracy

$$\sigma_x = \sigma_y = ((\sigma_{map}^2 + \sigma_{O_{new}}^2 + \sigma_{meas}^2)/2)^{1/2} = 1.6 \text{ m}$$

The expected accuracy is an combination of accuracy of the map ($\sigma_{\text{map}} = (\sigma_{\text{Xmap}}^2 + \sigma_{\text{Ymap}}^2)^{1/2} \cong 1.4$ m), orthoimage ($\sigma_{\text{O_new}} = 1.2$ m) and manual measurement of road crosses ($\sigma_{\text{meas}} = 1.5\text{pel} \approx 0.9$ m). The results are summarised in Tab. 2.6.

	New orthoimage O_LSM4		Existing orthoimage	
	X [m] / [pel]	Y [m] / [pel]	X [m] / [pel]	Y [m] / [pel]
Mean	0.06 / 0.1	-0.06 / -0.1	0.05 / 0.1	0.03 / 0.0
RMSE	1.14 / 1.8	1.24 / 2.0	1.01 / 1.6	1.11 / 1.8

Tab. 2.6 Root mean square errors and mean differences between the new produced orthoimage O_LSM4 and the topographic map. The results of comparison of the existing orthoimage and the map are added to illustrate a slight decrease in quality in the new orthoimage.

It can be concluded that RMSE values correspond to the expected accuracy. Fig. 2.11 gives an overview of the direction and magnitude of the error vectors.

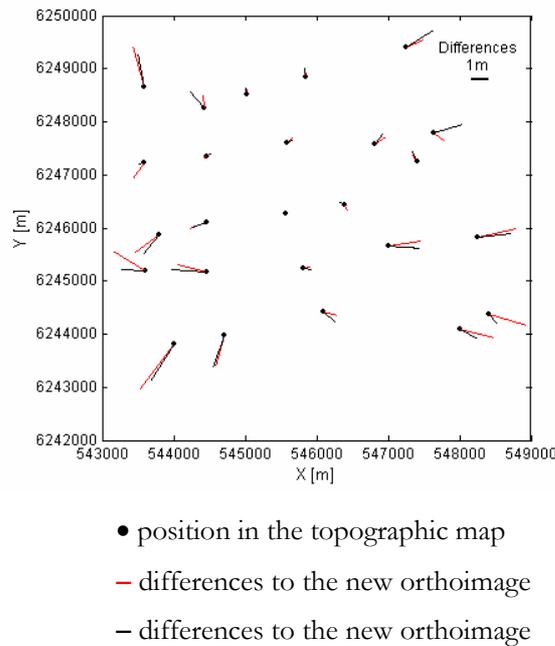


Fig. 2.11 Differences between TOP10Dk and the new (red line) and existing orthoimage (black line) measured in 25 road crosses.

Differences between the new orthoimage and the map are somewhat bigger than the differences measured in the existing orthoimage, especially at the edges of the image. The differences between the existing data sets can be explained by differences in orientation data of images used for derivation of both products and mapping accuracy. The reason for increasing the vectors in the new orthoimage is again in height differences between the DTM

used for orthoimage derivation and the topographic map. The radial shift dR between the position in the new orthoimage and the map can be calculated according to formula 2.4. Fig. 2.12 shows geometric relations used for derivation of the formula.

$$dR = dR_1 + dR_2 \tag{2.4}$$

$$dR_1 = \frac{R}{H} dh_1$$

$$dR_2 = \frac{R + dR_1}{H} dh_2$$

H height of the projective centre above the point in the topographic map

R radial distance between the nadir (N) and the point in the topographic map

dh_1 .. height difference between the map and the DTM used for derivation of an existing orthoimage

dh_2 .. height difference between the map and the DTM used for derivation of a new orthoimage

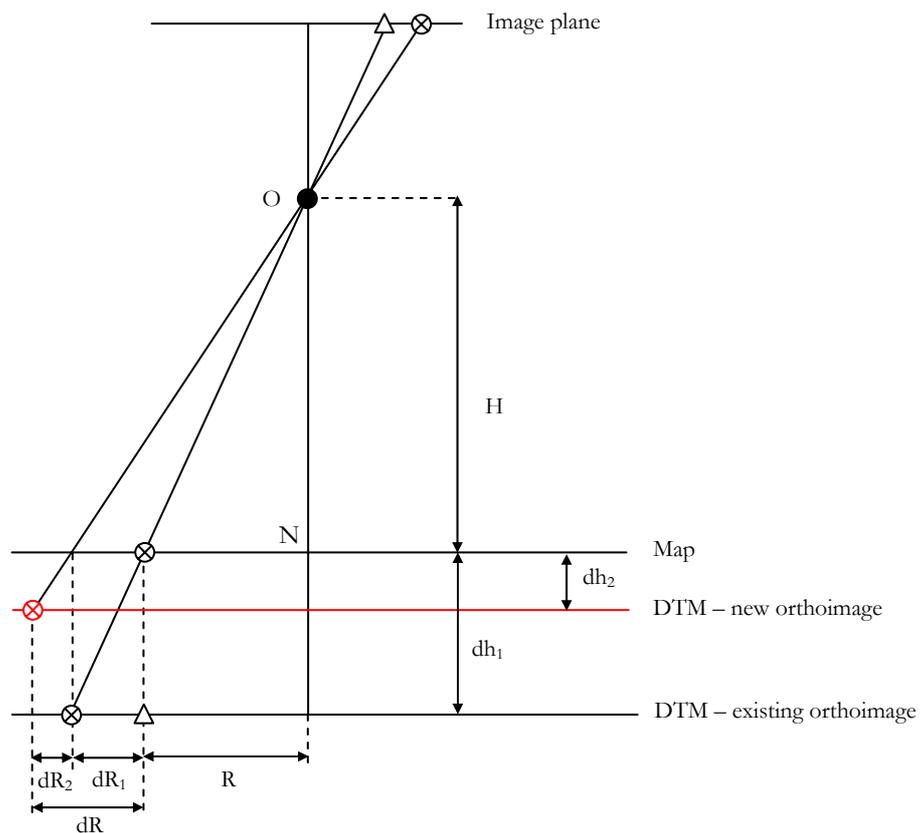


Fig. 2.12 Radial displacement of a road crossing (\otimes) caused by height differences of used data sets. An object in the existing orthoimage with co-ordinates corresponding to the road crossing in the map is marked with a triangle.

In the given data set the same DTM was used for deriving the old and new orthoimages. Therefore the smallest radial shift between positions of corresponding points in the map and the new orthoimage dR corresponds to the shift dR_1 between the existing orthoimage and the map and can be achieved only when the heights of control points are taken from DTM. Then the red vectors in Fig. 2.8 are identical with the black ones (within the measuring accuracy). When the map data were used for deriving orientation parameters, the increment dR_2 of the vectors appears. Its size changes locally depending on agreement between DTM and map heights and mapping accuracy. For average difference between TOP10DK and given DTM of 0.6 m, $dR_2 \cong 0.6$ m at the edge of the image which is a value corresponding to the accuracy of measurement. Such a value is not very noticeable when the road crossings are measured manually because it is in the level of the measuring accuracy.

In case that for derivation of the new orthoimage a new DTM exists and corresponds to the topographic map, the shift $dR=dR_1$ will be still present in the new orthoimage regardless which heights are used for orientation. If there is a good agreement between DTMs used for the orthoimage generation and the topographic map but some other systematic errors are present in the existing orthoimage (rotations, shifts in the position of the projective centre) they will appear in the new orthoimage. In order to guarantee a good agreement between the map and the new orthoimage feature based matching or area based matching with a preliminary step of vector to raster conversion must be applied.

2.6 Evaluation of the method and of the results

Evaluation of the applied method is done according to six points stated as the goals of this investigation in chapter 2.3.1. At the same time some suggestions for improvements or a more general solution are given.

1) **Selection of control information.** Developed programs carried out both finding the position of control points i.e. objects of given attributes in the ASCII presentation of the topographic map and the derivation of patches from image data. The only step missing for achievement of full automation was the measurement of fiducial marks in the aerial image. This procedure is well described in literature (e.g. Heipke, 1997). The part missing among all developed programs in order to automate also this step is finding an approximate position of the fiducial marks in the image.

2) **Check of suitability of image patches for matching.** Regarding the available topographic data, the image patches in areas where a lot of mismatches could appear (in the given images only forests) were excluded manually. This task could be automated by further programming or questioning a database if available (see chapter 2.3.2). When using road crossings in area based matching, a geometric attribute, namely angle of road intersection should be checked. If roads intersect in a small angle, it can happen that a template does not contain enough structures (only linear object) which leads to mismatches. If well defined points as road crossings that are unique with respect to their surroundings are not chosen as control points, autocorrelation would be a practical tool to apply in order to find about a uniqueness of a chosen point and the correct size of the matching windows.

3) **Improvement of approximation of orientation parameters by hierarchical approach.** Applying an image and object pyramid was successful with respect to improving approximations of orientation parameters. A small amount of control points is necessary for orientation of images with reduced geometric resolution (8 –12 well-distributed points were sufficient, some outliers can appear). Their derivation and measurement as well as calculation of orientation parameters can be carried out with the same level of automation as for the image with the highest (original) geometric resolution. The matching by cross-correlation is sufficient for that purpose.

4) **Image matching with subpixel accuracy.** The importance of LSM with four or six parameters rises as soon as differences in rotations and scale between an orthoimage and a new aerial image exist. In the performed test, applying of methods with subpixel accuracy for the measurement of control points brought improvements in fit between the new and existing orthoimage of about 0.5 pel in both co-ordinates in comparison to calculations based on cross-correlation only. A much better improvement cannot be expected also due to time changes and different illumination of the scene. A pre-requisite of a good agreement in all input data must be fulfilled.

5) **Detection of outliers.** After matching of selected control points by means of cross-correlation, the values of correlation coefficient were checked against a threshold of 0.5. The threshold was set rather low in order to eliminate only the biggest possible blunders and minimum of correct points. In the next step robust adjustment was applied. A combination of these two approaches seems to be successful when the number of mismatches is minimised by

prior selecting control points that are suitable for area based matching. In case of random points, more measurements are necessary and an application of thresholding of similarity measures gets a higher meaning in order to improve the ratio between erroneous and correct measurements to the level that is possible to handle with robust adjustment. When excluding measurements, a distribution of control points in the image must be watched. Therefore the centre of gravity and the area covered by control points that were not weighted down during adjustment were calculated.

6) **Quality control of the new orthoimage.** The comparison of the new produced orthoimage with the existing orthoimage was carried out automatically by means of LSM with four parameters. The importance of a good geometric agreement in the input data sets was shown. The comparison with the topographic map was carried out manually. This step should be automated in order to avoid a low accuracy manual detection of corresponding points.

The results showed that a fit of 0.5 pel in both X and Y co-ordinates between the existing and the new generated orthoimage which orientation parameters were derived by means of presented method can be achieved. Relatively high number of control points (about 60 per image) and least squares matching with four geometrical parameters allowing for scaling and rotation of image patches were used in order to accomplish such a result. Except of the measurement of fiducial marks and excluding points from areas inapplicable for area based matching, all steps in the process were carried out automatically.

The new orthoimage has errors that are present in the old orthoimage. They are caused by erroneous DTM or orientation parameters. If only relative relations are of interest, the new orthoimage can be used for change detection and map updating. If absolute accuracy should be improved, an orientation by means of aerotriangulation with well defined control points and GPS-IMU measurements must be used for the derivation of orientation parameters of the new image.

As soon as orthoimages and DTMs with sufficient accuracy are established, the presented method can be of a high value both for orthoimage production and map updating due to its accuracy and low costs. The database of georeferenced image patches of time invariant objects can then be established and maintained. It would save several steps in the process of image orientation. An object pyramid should be created in order to be able to start with rough approximations of orientation parameters.

3. Automatic DTM check and correction based on orthoimages

Mathematical models of landscape that provide information about a spatial position of points on the earth surface and optionally natural or man made structures raised above it have several important applications. Planning of road constructions, intervisibility for telecommunication purposes, hydrologic studies and orthoimage production are some of the examples. Each of the applications requires a different accuracy of the model. The mapping agencies and private companies dealing with orthoimage production have relatively high demands for the quality of such models. First, they should be free from systematic errors and outliers. Second, the overall accuracy should be sufficient to assure a required planimetric accuracy of the well-defined points in the orthoimage (see formula 2.1). Methods for checking and possibly correcting the models automatically are therefore demanded.

The main concern of this chapter is to investigate and improve a method of checking and correcting an existing digital terrain model (DTM) that is used for the derivation of orthoimages from medium scale aerial images. First, a definition of the different types of the digital height models is given. The basic characteristics, as well as the methods of acquisition of digital terrain models are discussed. Second, basics of the automatic derivation of the DTMs from aerial images are described. The last part deals with a method of checking and correcting the DTMs based on finding conjugate points in two overlapping orthoimages. The results of practical tests of the method using a DTM automatically derived from images at the scale of 1:25 000 are presented.

3.1 Digital terrain models and their quality

Real landscapes are usually too complex for an exact analytical modeling. Therefore the information is mostly made of samples. A **Digital Elevation Model** (DEM) can be defined as a numerical model of an object surface that together with a given mathematical interpolation method enables to calculate a height of any point on the object surface covered by the model (Jacobi, 1994). DHM is a general term. Concerning the earth surface, it can carry information about the ground elevation. Then the model is called a **Digital Terrain Model** (DTM). On the other hand, a **Digital Surface Model** (DSM) contains the highest elevation at each point (e.g. houses or canopy). Because the DTMs are the main concern of this and following chapters, attention will be paid neither to DEMs nor to DSMs further although many of properties that will be mentioned are common to all models.

DTM is usually represented by a set of points with X, Y, and Z co-ordinates in a given reference system. The higher the density of the points is, the better the model describes the real landscape. The specifications of DTM contain information about (Kasser, 2002):

- Datum (ellipsoid, elevation origin), map projection, height definition, metric units
- Geographic location (corners of an area covered by the model)
- Grid structure

The grid structure can be irregular (a triangular irregular network (TIN) or digitized contour lines) or regular (usually a form of a square mesh regular grid). Fig. 3.1 shows examples of different grid structures.



Fig. 3.1 DTM representation by a square mesh regular grid (left), contour lines (middle) and TIN (right). Break lines in TIN are marked with red color.

The main advantage of a TIN is that the characteristic lines and points describing morphologic properties of the terrain as break lines, main crests, summits, etc. are a part of the model. TIN is usually saved in a form of a table containing information about which points create a triangle and which the three neighboring triangles are. The model has quite high requirements for the storage capacity. The size of the square mesh is the main characteristics of the regular grid representation of the DTM. The regular grid has an advantage of a fast search and interpolation within the grid. On the other hand some of the morphologic structures are not modeled properly at a given size of the mesh and therefore can disappear or become deformed. This problem can be solved by creating a DTM which does not have a constant size of the square mesh and which is acquired by progressive sampling.

It should be mentioned that a DTM describes an elevation of a point as a function of its position $Z=f(X, Y)$, e.g. for a given position only one height can be calculated. This is a main difference to 3D models created e.g. in CAD systems. Therefore a term 2 ½ dimensional

model is usually used with respect to DTMs (Kasser, 2001, Kraus, 1997). Depending on the grid, the Z value for a point in a general position X and Y is calculated by means of a suitable interpolation method e.g. a linear interpolation for TIN and bilinear interpolation for a square mesh regular grid (see also chapter 1.1).

There are several methods for the acquisition of DTM data:

- a) **Digitizing contour lines** from existing topographic maps. The height information about the terrain in the form of contour lines is a part of the topographic maps in many countries with a land surveying and mapping tradition. The contour lines were derived and in areas of frequent changes updated by means of new terrestrial and photogrammetric measurements. The demand of quick creation of nation-wide DTMs caused that the existing maps were scanned and contour lines digitized instead of new mapping. The consequence is that the height information is several decades old in many areas.
- b) **Ground surveys** by means of tacheometry or GPS may be accurate (σ_z is in the range from 0.15 m to 0.02 m) but suitable only for small or complicated areas due to low speed and high costs.
- c) **Aerial photogrammetry** is the main method for creating DTMs due to its accuracy ($\sigma_z \leq 0.02\%$ is achievable), speed and therefore relatively low costs. The method has a limited use or cannot be used at all in forest, water, snow or sand areas. Automatic methods for DTM derivation from aerial images are discussed in chapter 3.2.
- d) **Airborne laserscanning** is a relatively new method that in comparison to photogrammetry can be used both day and night under good weather conditions (without low clouds, rain, fog). It is suitable for measuring in forest, sand, water or snow areas and also cities. Its typical accuracy is between 0.15m – 0.20m in height and 0.3m – 1.0 m in position and the density of points can be down to 1 point per 1m². In comparison to photogrammetry, its disadvantage is relatively low speed of measurement.
- e) **Interferometric radar and satellite images** are another sources for DTM acquisition but due to their lower accuracy ($\sigma_z > 5$ m) are not discussed further.

The quality of a DTM is usually evaluated both by an accuracy of a single point interpolated from a DTM but also by the capability of the model to describe the real terrain. Furthermore, criteria of completeness, reliability, consistency, and uniformity of the accuracy distribution within DTM must be fulfilled (Ackermann, 1996b). In general, parameters that have influence on the accuracy of a single point which height was derived from the DTM are:

a) **Accuracy of the measurement of DTM points.** It depends on the method used for DTM acquisition and was discussed in the upper paragraph. Systematic and outliers in the heights are especially of concern.

b) **Grid density and structure.** Only details with the distance larger than two times a grid size can be registered. Reducing the density of points makes the terrain model smoother (see Fig. 3.2).

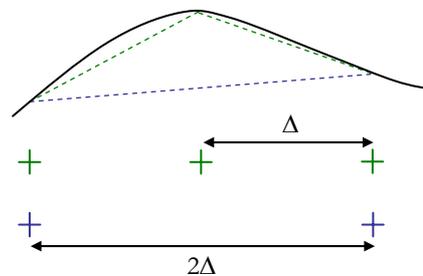


Fig. 3.2 Influence of the size of the mesh on the accuracy of the modeled terrain. Grid points with the distance of Δ model the terrain undulation relatively well while modelling with the mesh distance of 2Δ smoothes the terrain.

The density of DTM points is chosen according to the terrain type (steep and rolling terrain requires higher density of points than a flat one) and according to accuracy requirements. As mentioned before, TINs are better suited for describing morphological variations of the terrain. Breaklines and all other significant terrain points have to be added to the regular grids in order to avoid outliers.

c) **Interpolation method** used for calculating grid points from observed points. As pointed out in (Ackermann, 1996b), as long as dense and high quality data are available, the different interpolation methods perform similar. But as soon as the point distribution is poor, errors can occur. Therefore, interpolation programs should have tools for self-diagnostic and inform the user about areas of possible weak DTM quality.

DTMs derived from contour lines as well as by methods b)-e) contain outliers and/or systematic errors (Larsen, 1998, Wind, 2001). The quality check of DTMs is therefore necessary. Several manual approaches exist:

- Measurement of single points and profiles both by terrestrial methods or by stereophotogrammetric methods
- Visual inspection by
 - superimposition of a DTM grid or of contour lines onto a stereoscopic model (even small errors can be found)

- displaying overlapping orthoimages derived from the tested DTM as a stereoscopic model (areas with wrong heights will appear as hills or valleys in a flat landscape, for theoretical explanation see chapter 3.2.1)
- 3D perspective views or DTM rendering (for discovering outliers)

Manual approaches are time consuming in case of DTMs covering a large area. Automated procedures have therefore been suggested. They include both quality assessment during automatic DTM generation (Ackermann, 1996b, Gooch and Chandler, 2000) or are based on matching orthoimages derived from a stereopair and the tested DTM (see chapter 3.3).

3.2 Basic concepts of automatic DTM generation from aerial images

Nowadays, automatic derivation of DTMs from stereomodels is a standard tool of professional software packages as ‘Automated terrain extraction’ of Leica, ‘Match-T’ of Inpho, or ‘ImageStation automatic elevation’ of Z/I imaging. Although there are differences in the algorithms used in different packages, the general strategy consists of three steps:

- finding conjugate points (image matching)
- interpolating grid points (surface fitting)
- internal quality control

Finding conjugate points is done by feature and area based matching (cross-correlation and LSM). A hierarchical approach is applied. In the first approximation the terrain is considered as a horizontal plane. Orientation parameters of images used for the DTM derivation are known. Therefore epipolar geometry is used for finding conjugate interest points. The density of interest points is high. A so called parallax bound is used in order to exclude points on objects rising from the ground (house, trees, etc.). It must be set carefully with respect to the character of the mapped landscape (small value for flat areas and larger value for hilly or mountainous areas). Thresholds for similarity measures (w and q values of the Förstner operator, correlation coefficient, standard deviation of shifts derived by LSM) can be fixed or set as adaptive, i.e. some measurements are obtained also in poorly textured areas.

The points found by image matching are not evenly distributed. A DTM is usually required in a form of a regular (mostly rectangular) grid. Therefore the **interpolation** of elevations of grid points must be carried out. Different methods can be used such as polynomial or spline methods, finite elements, etc. (Schenk, 1996, Ackermann, 1996). As mentioned in the previous chapter, the differences between methods appear especially at the areas with a low density of points. In order to create a reliable DTM, identification of outliers on statistical bases and

integration of manually measured break lines should be included in the interpolation algorithms. The density of the grid points should correspond to the character of the terrain. An adaptive grid (denser in areas with slopes and sparser in flat areas) is an option for optimizing the amount of data. At the same time a DTM describing the terrain with a sufficient amount of detail is obtained.

During DTM generation an **internal quality control** can be carried out. The standard deviation of the calculated height can be derived. It is influenced by parameters such as number of interest points from which the height was calculated, distance to the interest points and their accuracy. In this way each of the DTM points receives a quality attribute. A thematic map based on point positions and attribute values gives a quick overview over the DTM quality and can be used as guidance for a DTM check and correction (see Fig. 3.3).

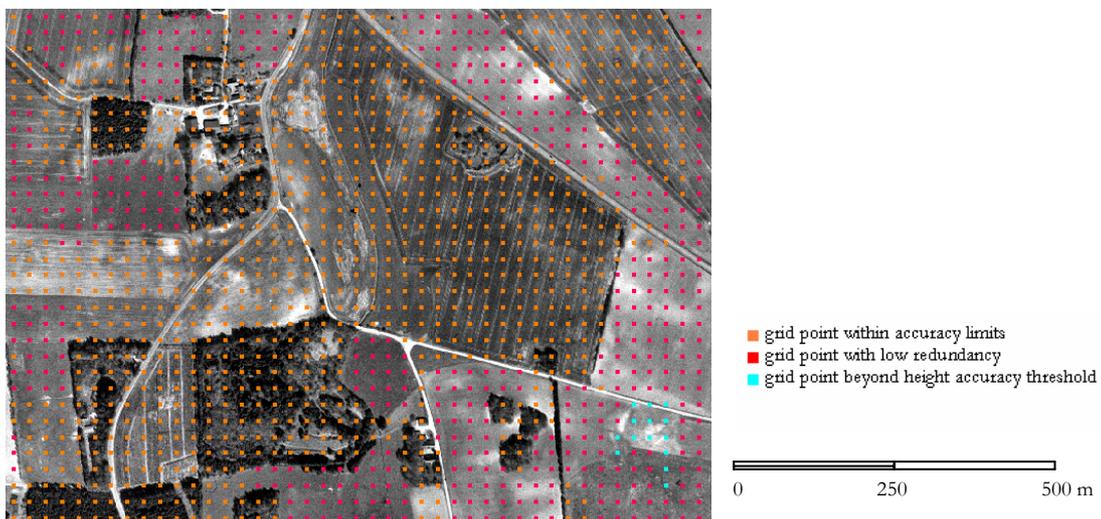


Fig. 3.3 DTM quality evaluation based on statistical values calculated during automated DTM derivation in ImageStation Automatic Elevation program of Z/I Imaging.

The accuracy of DTMs obtained by automatic measurements found in literature differs. Results published in (Ackermann and Schneider, 1992) showed $\sigma_z=0.005\%h$ (0.009%h) for images 1:22 600, image resolution of 15 μm (30 μm) and smooth hilly terrain. In mountainous terrain the values rose to $\sigma_z=0.011\%h$ (0.016%h). Results published in (Wind, 2001) revealed $\sigma_z=0.07\%h$ for images of a similar scale (1:25 000) and resolution (30 μm) in the flat or slightly rolling terrain. After excluding outliers and subtracting a systematic shift in elevation, an accuracy of $\sigma_z=0.017\%h$ was achieved. In both cases the reference data used for the

derivation of standard deviation values were measured manually in an analytical plotter. The test in the second mentioned publication was carried out for finding a suitable method for DTM derivation with a national coverage. The requirement from the Danish mapping agency was to achieve an accuracy of 0.5 m, e.g. 0.013‰ for images at the scale 1:25 000. Such an accuracy was achieved for images 1:15 000, but again after manual removing outliers. It leads to a conclusion that in spite of internal quality control of DTMs outliers can appear and therefore other check methods have to be applied.

3.3 Check of DTMs by means of overlapping orthoimages

The outliers and/or systematic errors in DTMs derived from contour lines or by automatic methods based on aerial images or from laserscanning data exist. Manual checking by methods mentioned in the chapter 3.1 is time consuming especially if DTMs cover large areas. The need of automatic methods giving a quick overview about the quality of a DTM was stated as important both by state and private mapping agencies.

An error detecting technique for automatically generated DTM is suggested in (Gooch and Chandler, 2000). A system called 'Failure Warning Model' is described. The DTM is automatically derived twice with a slightly different parameter setting. Finding weak areas of the terrain model is based on an idea that at areas with low accuracy or reliability differences in two models will appear. The method is promising and can be easily integrated into standard photogrammetric software packages for an automatic DTM generation. Nevertheless, it does not solve the problem of checking an existing DTM derived from digitized contour lines.

The method that gives a solution to above mentioned problem was suggested in (Norvelle, 1996, Li et al. 1996). A residual horizontal parallax of conjugate points in two overlapping orthoimages is found by image matching. Since orthoimages should not contain any relief displacement, the parallaxes are an indication of elevation errors. Because the method can fail at areas of low or repetitive texture, a priori detection and exclusion of such areas was proposed in (Krupnik, 1998). The approach of finding errors in DTM by comparison of two orthoimages has been found interesting and has become an object of further investigations that are described in the following chapter.

3.3.1 Principles of the method

Assuming both a DTM and orientation of an image stereopair free of errors, shifts between corresponding points in two overlapping orthoimages should not appear. If errors in the DTM exist, horizontal parallaxes dX in conjugate points can be measured and corresponding height differences dh can be calculated (compare Fig. 3.4 and formula 3.1). The derivation of the formula and first experiments with the method are presented in (Höhle and Potůčková, 2004).

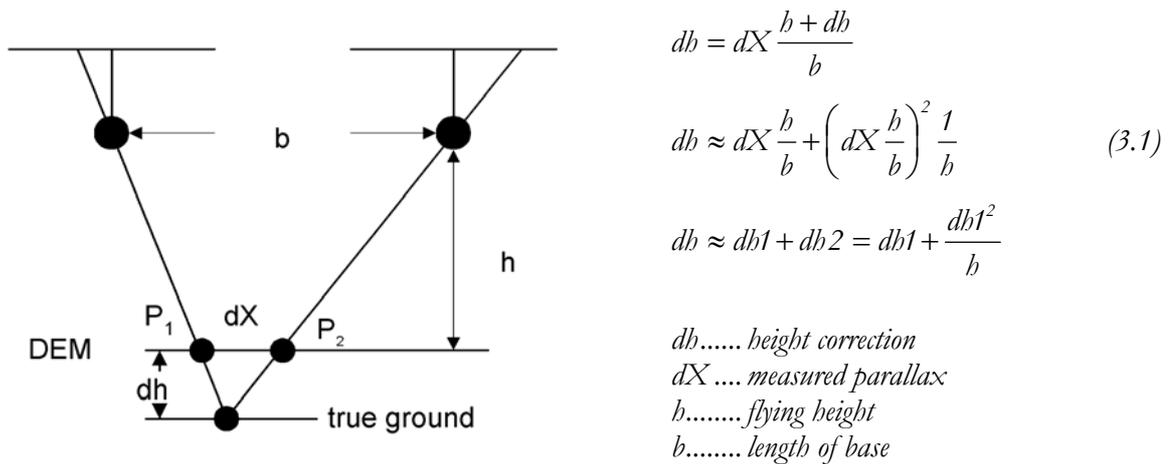


Fig. 3.4 Principle of the method for checking and correcting DTMs based on two overlapping orthoimages.

The parallaxes dX can be found by feature or area based matching. The formula 3.1 is approximate and therefore the calculation has an iterative character, i.e. the DTM is corrected of values dh , new orthoimages are derived and the whole procedure can be repeated. The number of iterations depends on the size of DTM error. The second term in formula 3.4 has an importance only for a lower flying height and larger height differences as Fig. 3.5 shows.

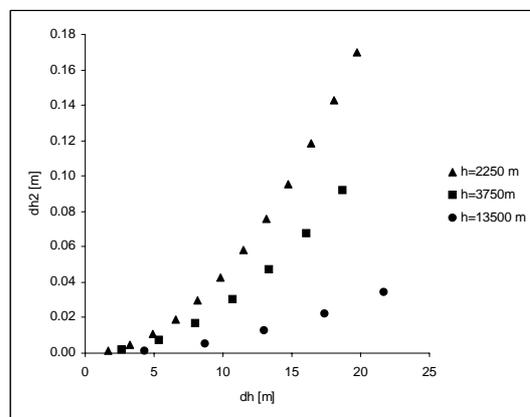


Fig. 3.5 Influence of the term $db2$ on the value of height difference $dh=db1+db2$. The calculation was done for image scales 1:15 000, 1:25 000, 1:90 000, a wide-angle camera ($c=15$ cm), and overlap 60%.

3.3.2 Test of the method

The method of the DTM check that is based on determining parallaxes in overlapping orthoimages was tested on the data set described in the first part of this chapter. The developed calculation strategy is discussed and the results are evaluated afterwards.

3.3.2.1 Data set description

The data set covers an area around of the town of Gistrup, Denmark. It consists of three stereopairs of two different scales and an automatically derived DTM. All the data including orientation parameters for images were taken from a Ph.D. project on evaluating performance of software packages for automated DTM generation (Wind, 1996). All the calculations were done in the Danish national reference system S34J and a vertical datum DNN. The following data were used:

Aerial images:

- a) one image stereopair, scale 1:25 000, scanning resolution 30 μm
- b) two image stereopairs, scale 1:15 000, scanning resolution 15 μm

The black & white images were taken in May 1996 by a wide-angle aerial camera RMK TOP 15. All the measurements in images necessary for computing orientation parameters were done in an analytical stereoplotter (Zeiss Planicomp C120). Control points were measured by kinematic GPS with an accuracy of $\sigma_x = \sigma_y \approx 1\text{cm}$, $\sigma_z = 1.5\text{cm}$. The images of both scales were part of larger blocks. The orientation parameters were calculated by means of a bundle block adjustment (software package Bingo). Tab. 3.1 contains standard deviations of obtained orientation parameters of images in both scales.

	σ_{X_0} [m]	σ_{Y_0} [m]	σ_{Z_0} [m]	σ_{ω} [mgon]	σ_{ϕ} [mgon]	σ_{κ} [mgon]
1: 25 000	0.250	0.224	0.146	4.0	3.2	1.4
1: 15 000	0.167	0.189	0.101	4.3	4.6	1.9

Tab. 3.1 Mean standard deviations of orientation parameters of images 1:25 000 and 1:15 000 obtained in bundle block adjustment

Digital terrain model

The DTM was derived automatically from a stereopair 1:25 000 by means of the software package 'MatchT' of Inpho. It has a form of a square mesh regular grid with the mesh size of 25 m. It is an original DTM without any manual editing or outlier elimination. According to

literature referring to ‘MatchT’ program (Ackerman, 1992), it should be possible to achieve an accuracy of $\sigma_z=0.01\%h$, i.e. 0.38 m. The results published in (Karras et al., 1998 and Wind, 2001) were slightly worse, about 0.014 - 0.017%h and were achieved after DTM editing and excluding outliers. Therefore, an accuracy of $\sigma_z = 0.02\%h$, i.e. 0.75 m was taken as realistic for the evaluation of the test model.

Reference data

In order to evaluate the quality of the DTM and the performance of the check method, reference heights were measured in images 1:5000 in the analytical stereoplotter by an experienced operator. The accuracy of the reference heights is $\sigma_z=0.1m$. The reference data do not cover all models but only some parts (see Fig. 3.6). All the reference heights were measured on the ground in a grid of 25 m.



Fig. 3.6 Overview of the model covered by images 1:25 000. The white polygons mark areas where reference data were available. Two models 1:15 000 covered the areas of the two rectangles in the left part of the figure.

Orthoimages

Orthoimages were derived using the software package ‘ImageStation Base Rectifier’ of Z/I Imaging. Depending on the image scale and the scanning resolution, the ground sampling distance (gsd) was set to 0.75 m and 0.225 m respectively. In order to minimize the interpolation errors, a distance between anchor points of 2 pixels and bilinear interpolation were set as parameters for the orthoimage derivation.

The area of the models is mostly open land area. Forest and urban areas cover a relatively small part of the model 1:25 000. The landscape is relatively flat or mildly rolling. Fig. 3.6 gives an overview of the entire area of the model 1:25 000. The borders of the areas where the reference heights were measured and the DTM tested are marked.

Comparison of the DTM and reference data

The DTM heights were compared with reference heights at almost 6000 points. Areas that would lead to confusions with respect to the tested method were excluded from the comparison and all further calculations, for example the dense forest area in the southeastern part of the model 1:25 000 (compare Fig. 3.6 and 3.7). Fig. 3.7 shows distribution of errors over the test area and Tab. 3.2 summarizes results of the comparison.

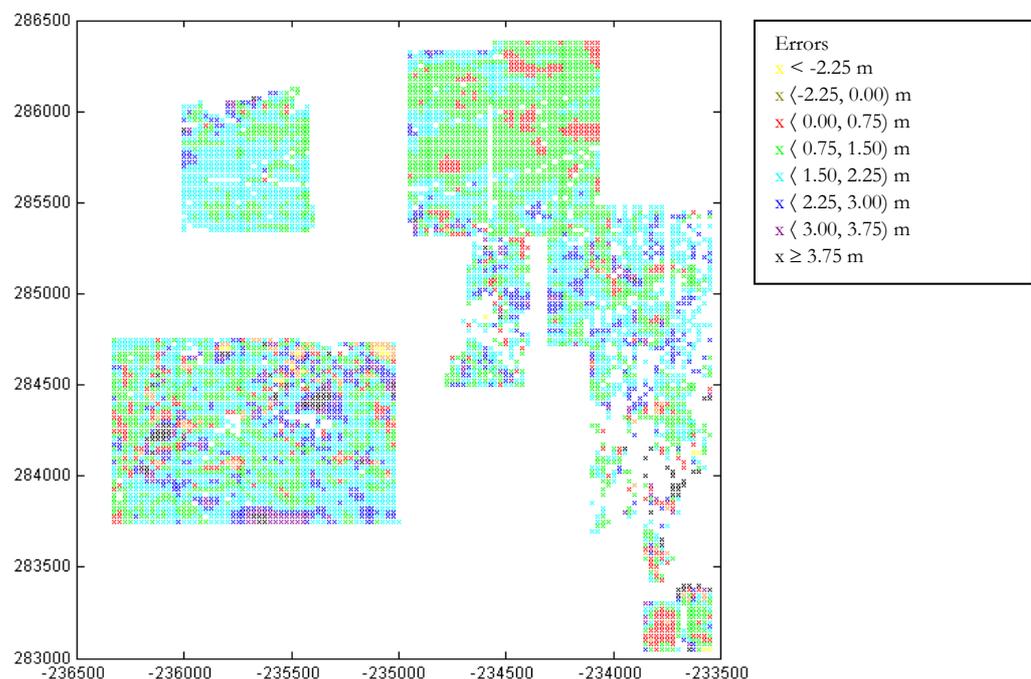


Fig. 3.7 Distribution of errors in an automatically generated DTM. Statistical characteristics can be found in Tab.3.2. The points with the same differences to the reference data appear usually in groups. The points are missing at dense forest and built-up areas.

It is obvious from both Tab. 3.2 and Fig. 3.7 that the expected accuracy of $\sigma_z = 0.02\%h=0.75\text{ m}$ was not achieved. The $RMSE=1.85\text{ m}$ is higher with a factor of 2.5. A systematic shift of the DTM elevations exists. Even after its subtraction the standard deviation $\sigma_z=0.95\text{ m}$ is high. In following sections, the source of the systematic shift will be analyzed and suggestions for the automatic detection of the erroneous points (areas) will be given.

Number of points	5778
Max. error [m]	16.47
Min. error [m]	-5.52
Mean [m]	1.61
RMSE [m]	1.87
σ_z [m]	0.95
Number of outliers	14% (805 points)
Number of outliers after subtracting the mean	2% (117 points)

Tab. 3.2 Comparison of the DTM with the reference heights. A required accuracy of $\sigma = 0.75m$ is not achieved. An error $abs(h_{DTM}-h_{ref}) > 3\sigma$, e.g. 2.25 m is considered as an outlier.

3.3.2.3 Calculation strategy

As mentioned in the previous section, the grid points in areas with a complete cover of trees and buildings were excluded from the test because the found DTM corrections would correspond to the top of those objects and not to the ground. The calculation itself consists of four steps:

- Derivation of image patches from the left and the right orthoimages and their matching by means of area based methods (correlation coefficient and LSM with two geometrical parameters). The matching is carried out only along an epipolar line (see Fig. 3.8).
- Calculation of horizontal parallaxes and height corrections according to formula 3.1
- Decision in which areas it is necessary to correct the DTM
- Evaluation of step c) by a comparison with reference data

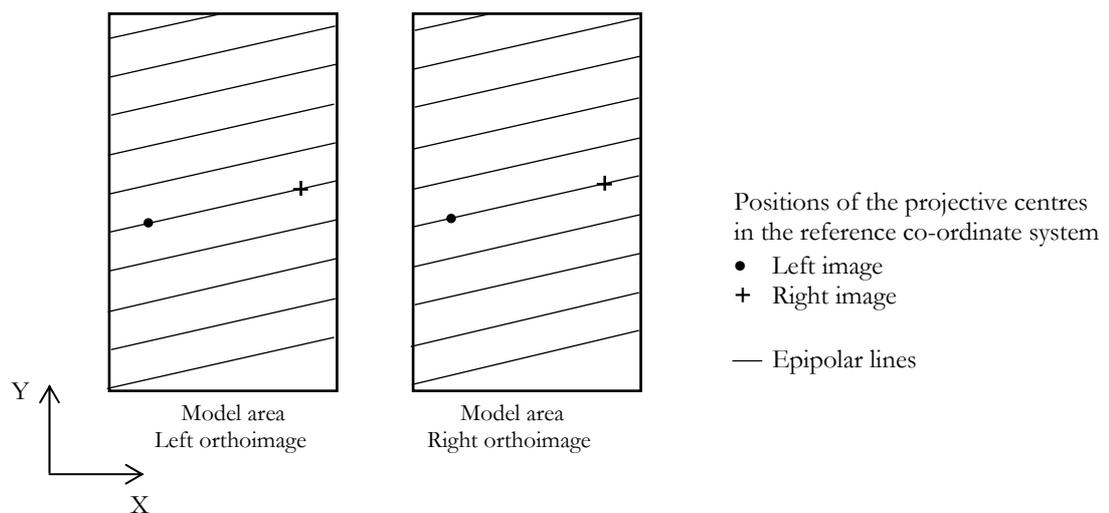


Fig. 3.8 Epipolar lines in two overlapping orthoimages. Epipolar lines are parallel to the base, e.g. a line connecting the perspective centres of original aerial images.

The whole procedure was repeated twice. First, the matching was carried out directly in grid points. Second, points in a close surrounding of a grid point were matched in order to get more reliable results. After obtaining results of area based matching and after calculating the height corrections, the DTM points were divided into two categories. The first category contained points that did not have to be checked or corrected and the second category contained points where the corrections are high or the method failed. The decision was based on:

- thresholding similarity measures
- thresholding the size of correction
- number of points with the same height in a close neighborhood of a grid point

In the following text, the points that do not need any correction are called ‘accepted’ or marked with G (stands for ‘green’ according to the traffic light principle). The points that need further investigation are marked with R (for ‘red’).

3.3.2.3 Results – Imagery 1:25 000

Large parts of the test area are open fields with low texture. Therefore, the size of the templates is chosen rather large $r_T \times c_T = 21 \text{ pel} \times 21 \text{ pel}^2$ (or $15.75 \times 15.75 \text{ m}^2$). The number of columns in the search area c_S must enable finding parallaxes corresponding to the highest errors in heights (see Fig. 3.9):

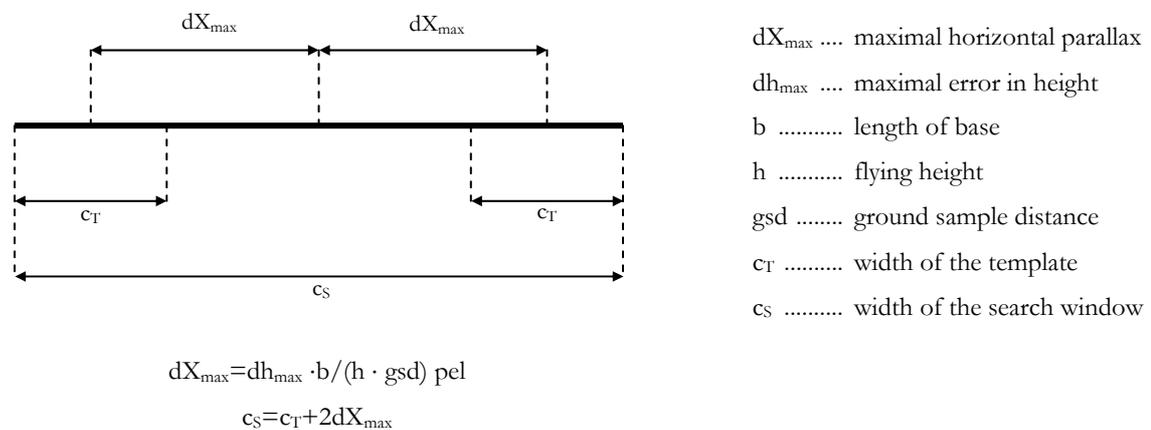


Fig. 3.9 Derivation of the width of the search window c_S . It depends on the value of the maximal horizontal parallax dX_{\max} that can appear and the size of the template c_T .

The horizontal parallax of $dX_{\max} = 14 \text{ pel}$ corresponds to the maximal height difference $dh_{\max} = 16.5 \text{ m}$ (see Tab. 3.2). Therefore the width of the search window should be 49 pel. One pixel was added to both sides in order to cover possible errors from orientation. Thus, search areas with the width of $c_S = 51 \text{ pel}$ was used for calculations. In the given test data, the flying

direction was almost parallel to the X-axis of the reference co-ordinate system. Epipolar lines are therefore parallel to rows of derived orthoimages. The matching was done along an epipolar line ± 1 pel in order to take care of possible orientation errors. Thus, the number of rows of the search area $r_s = r_r + 2 \text{ pel} = 23 \text{ pel}$.

The DTM corrections were calculated and added to the DTM heights. In order to see if any improvement of the DTM was achieved, the corrected DTM heights and reference heights were compared. After the traffic light principle, the points were divided into three categories:

G_{ref} : The absolute value of height error after correction is less than 3σ ($\sigma=0.75 \text{ m}$)

Y_{ref} : The absolute value of height error after correction is higher than 3σ but in comparison with an original DTM an improvement was achieved. The next iteration of the calculation of corrections could bring further improvement.

R_{ref} : The absolute value of height error is after correction higher than 3σ , e.g. 2.25 m

The results are summarized in Tab. 3.3.

	G_{ref}		Y_{ref}		R_{ref}	
	Original DTM	Corrected DTM	Original DTM	Corrected DTM	Original DTM	Corrected DTM
No. of points	4697 (81%)		87 (2%)		994 (17%)	
Min. error [m]	-5.52	-2.14	1.07	-2.99	-2.86	-17.39
Max. error [m]	12.50	2.25	16.47	4.59	12.40	23.82
Mean [m]	1.51	1.24	3.81	2.47	1.88	3.14
RMSE [m]	1.71	1.38	4.58	2.68	2.14	8.53
σ [m]	0.81	0.62	2.54	1.04	1.02	7.93

Tab. 3.3 Comparison of original DTM and corrected DTM with the reference data. The method worked successfully in 83% of the tested area ($G_{ref}+Y_{ref}$).

The results in Tab. 3.3 show that the systematic shift is present also after application of the correction method. It can be concluded that its origin is not in the process of the DTM derivation itself but most probably in the orientation data. Because the same orientation data was used for both the DTM and orthoimage derivation, it could not be corrected. The shift somewhat devaluates the results of the investigated correction method. Therefore in all the following text, the results obtained after subtracting the systematic shift of 1.61m (see Tab. 3.2) are presented. Tab. 3.4 shows the comparison of the corrected DTM with the reference data again but after subtracting the systematic shift.

	G _{ref}		Y _{ref}		R _{ref}	
	Original DTM	Corrected DTM	Original DTM	Corrected DTM	Original DTM	Corrected DTM
No. of points	5109 (88%)		3 (1%)		666 (11%)	
Min. error [m]	-7.13	-2.25	-2.99	-2.71	-2.46	-19.00
Max. error [m]	14.86	2.24	5.86	2.98	10.80	22.21
Mean [m]	-0.04	-0.17	2.88	-0.81	0.29	1.13
RMSE [m]	0.90	0.70	5.05	2.80	1.22	9.87
σ [m]	0.90	0.68	4.15	2.68	1.19	9.81

Tab. 3.4 Comparison of original DTM and corrected DTM with the reference data after subtracting a systematic shift of 1.61 m.

Tab. 3.4 shows that in 89% of points the method worked well – the original accuracy was kept at error-free points (88%) or it was improved (1%). The remaining 11% are mostly newly introduced outliers. If the procedure of the DTM check should be carried out automatically, the areas where the method fails must be identified and recommended for check by means of an alternative method, for example visual inspection.

In the following text, three methods for an automatic dividing the studied DTM into the G and R areas are described and tested. They are based on evaluation of:

- A) Corrections calculated directly in grid points
- B) Average of corrected heights of surrounding points
- C) Histogram of corrected heights of surrounding points

A) Corrections calculated directly in grid points

A relation between height errors in the corrected DTM and similarity measures, namely correlation coefficient, an accuracy of least squares matching and image distance has been studied. All relations are displayed in Fig. 3.10.

It can be concluded from the Fig. 3.10 that by setting thresholds for similarity measures the number of outliers can be considerably reduced but some of them remain. Due to the correlation between the three presented measures (especially correlation coefficient and image distance – see chapter 1.2.1.5), empirical thresholds suggested in Fig. 3.10 have performed almost the same – about 30% of points are moved to the category ‘method fails’ and remaining 70% contain about 5% of outliers. After applying all three measures, 63% of the area was accepted but 3% of outliers was still present. The highest errors appeared in areas of open fields with low texture and none or linear structures.

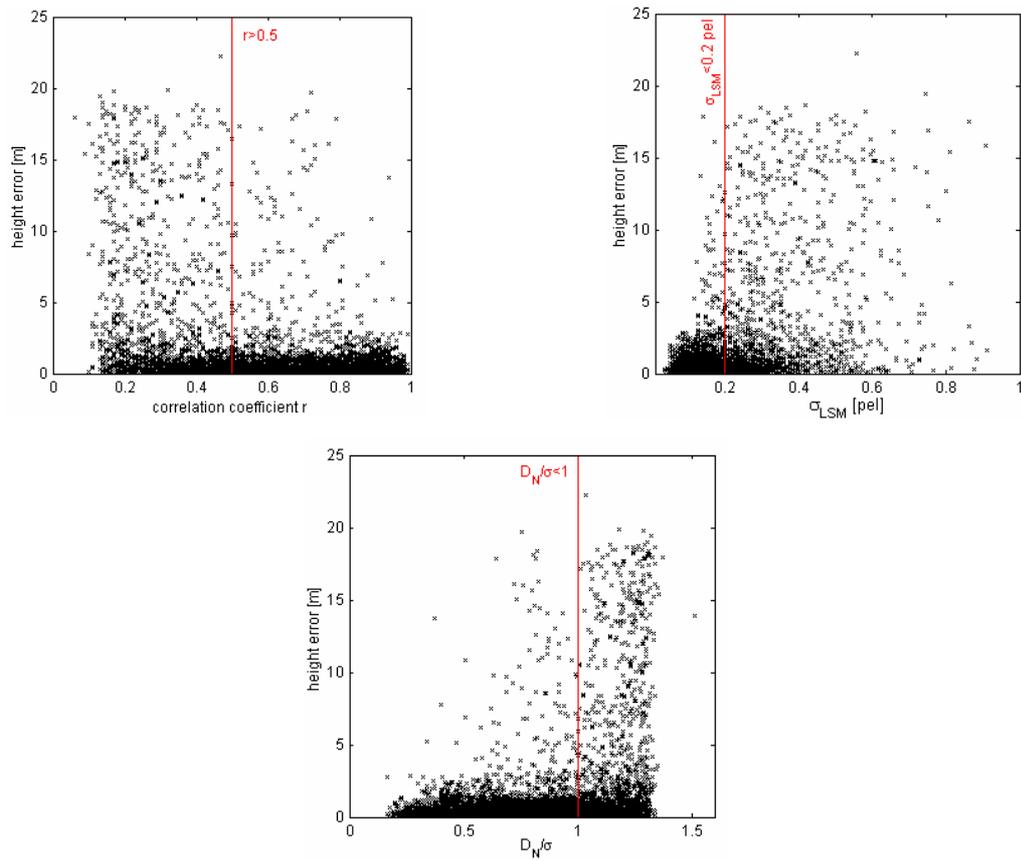


Fig. 3.10 Relation between height errors in the corrected DTM and attributes achieved during image matching procedure – correlation coefficient, standard deviation of shift parameters of LSM and a ratio between normalized distance and contrast. Thresholds (empirically set) for all attributes are marked with red lines.

The presented numbers show that the calculated corrections can be many times erroneous and their application can even decrease the DTM quality. This can be also concluded from the Fig. 3.11, which shows the relation between height errors and corrections. On the other hand, this figure also shows that if the corrections are small, the height errors are seldom large.

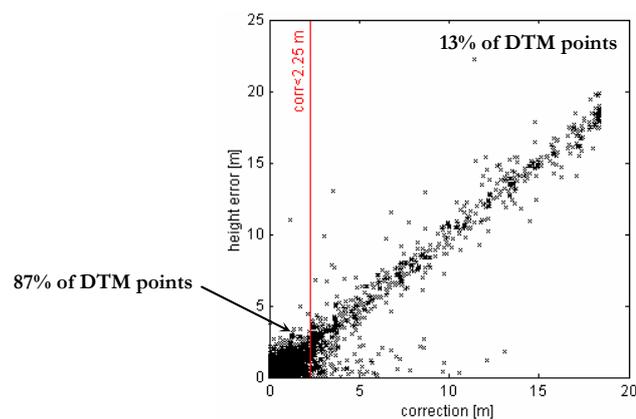


Fig. 3.11 The relation between the calculated height corrections and the height errors in the corrected DTM. The red line shows an empirical threshold for elimination of outliers. It is visible that most of outliers are introduced by the method itself.

Based on Fig. 3.11 the test area can be divided into two parts as mentioned before. The first one includes points which corrections are small and that is why the DTM is considered as correct and the second part consists of points where the corrections are somewhat larger because of the outliers in the DTM or failure of the method. Tab. 3.5 shows such a division for different correction thresholds. The thresholds were set as multiples of the expected accuracy $\sigma=0.75\text{m}$.

Correction threshold	T1= σ 0.75 m	T2= 2σ 1.50 m	T3= 3σ 2.25 m	T4= 4σ 3.00 m
Points accepted	3733 65%	4767 83%	5008 87%	5198 90%
Outliers included	7 0.2%	31 0.7%	54 1.1%	153 2.9%
RMSE [m] (%h)	0.53 (0.014)	0.63 (0.017)	0.67 (0.018)	0.70 (0.019)

Tab. 3.5 Dividing the entire DTM into two categories – ‘Accepted’ (G) and ‘For revision’ (R) based on the size of the derived correction. The table shows how many percent of all 5778 investigated points appear in the category ‘Accepted’ and how many outliers remain in that category. The RMSE value is calculated from the original DTM heights.

If thresholds T1 or T2 are applied, the number of outliers in the category ‘accepted’ points, i.e. points that will not be inspected further, is less than 1%. In order to guarantee results from matching, additional thresholds for the correlation coefficient and the standard deviation of least squares matching were included. Obtained results are summarized in Tab. 3.6.

From Tab. 3.6 can be concluded that the number of outliers in the area that is marked as without errors (G) decreases with an application of stronger criteria. By setting the three thresholds T1: corrections $<0.02\%h$, T_c: correlation coefficient >0.5 and T_{LSM}: $\sigma_{LSM} <0.2$ pel, 52% of the tested points were suggested for further investigation. The remaining 48% of points are assumed as correct. They contain only 0.1% of outliers which correspond well to the normal distribution. Moreover, the standard deviation calculated from those points revealed the value of 0.013%h that is also within the limit required by some national mapping agencies, e.g. the Danish National Survey and Cadastre requires 0.5m (Wind, 2001).

Thresholds		
correction [m]	T1= σ 0.75	T2= 2σ 1.50
T_r	0.5	0.5
σ_{LSM} [pel]	0.2	0.2
Points accepted	2755 48%	3379 58%
Outliers included	3 0.1%	16 0.5%
RMSE [m] %h	0.50 (0.013)	0.59 (0.016)
Max. error [m] (absolute value)	3.00	3.00

Tab. 3.6 Dividing the entire DTM into two categories – ‘Accepted’ (G) and ‘For revision’ (R) based on the size of the derived corrections and thresholds for correlation coefficient and accuracy of the shift parameters of LSM. The table shows how many percent of all 5778 investigated points appear in the category ‘Accepted’ (G) and how many outliers remain in that category when applying mentioned thresholds. The RMSE value is calculated from the original DTM heights.

B) Average of corrected heights of surrounding points

Other possibilities that would increase the reliability of results and the areas where the method works safely have been studied. The disadvantage of the procedure applied above is that a correction in a grid point is estimated only from one observation. A new procedure has therefore been designed. The area around a grid point is assumed to be a plane. The height corrections in several points in the close surrounding of the grid points are found and the corrected height of the grid point is derived. Because the test area is mostly flat and open land area, relatively large surrounding of $21 \times 21 \text{pel}^2$, i.e. $15.8 \times 15.8 \text{m}^2$ with the grid point in the middle was chosen. 25 even distributed templates of the size $15 \times 15 \text{pel}^2$ were matched per one grid point. The templates were extracted from the left orthoimage, the search areas $19 \times 43 \text{pel}^2$ from the right orthoimage. The conjugate points were again found by means of cross-correlation and LSM. The height corrections were calculated according to formula 3.1.

First, the corrected height of the grid point was calculated as an average of corrected heights of 25 surrounding points. At the same time thresholds for correlation coefficient $T_r > 0.3$ and accuracy of LSM $T_{LSM} < 0.2 \text{ pel}$ were applied in order to assure quality of matching in surrounding points. A comparison with reference data was carried out. The division of the

area was done again according to the traffic light principle and can be seen in Tab. 3.7. together with obtained differences to the reference data in the ‘green’ area.

	G_{ref}	Y_{ref}	R_{ref}
No. of points	4355 (75.4%)	8 (0.1%)	1415 (24.5%)

G_{ref}	Original DTM	Corrected DTM
Min. [m]	-6.10	-2.22
Max. [m]	14.86	2.24
Mean [m]	-0.03	-0.10
RMSE [m]	0.87	0.63
σ [m]	0.87	0.62

Tab. 3.7 Comparison of original DTM and corrected DTM with the reference data. The corrected DTM heights were obtained by averaging the corrected heights of 25 surrounding points. A threshold for correlation coefficient of 0.3 and standard deviation of shift parameters in LSM 0.2pel were also applied.

In order to make the decision about ‘accepted’ areas and areas that must be inspect by other check method automatically, the standard deviation of the calculated mean (σ_{mean}) of heights of surrounding points was derived and studied. Fig. 3.12. shows the histogram of σ_{mean} values for G and R areas from Tab. 3.7.

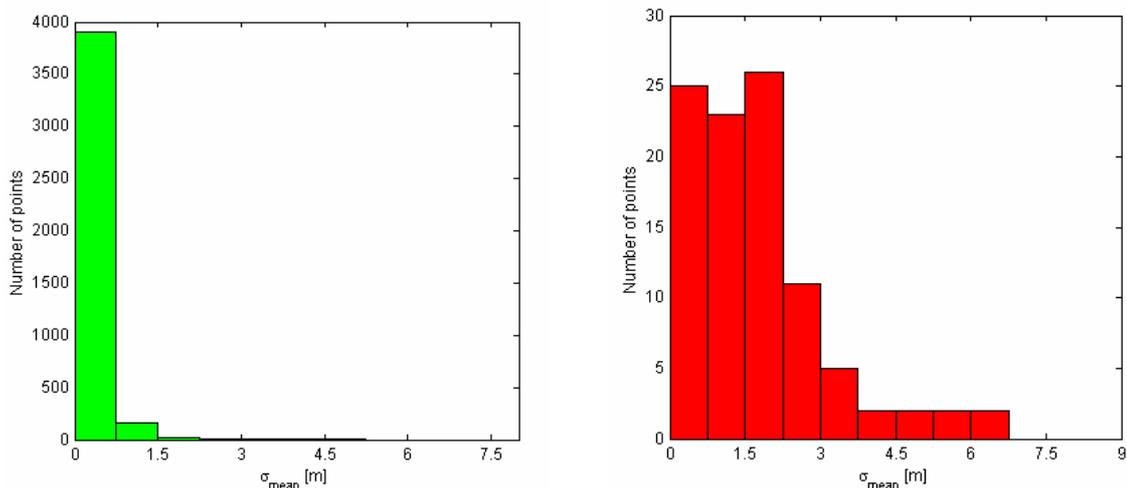


Fig. 3.12 Standard deviations of the mean height (σ_{mean}) calculated for points in the G and R areas. The division was made by comparison with the reference data.

Based on the histograms the value $\sigma_{mean} = 0.75m$, that also corresponds to the expected accuracy of the DTM, was used as a threshold. In order to minimize the number of outliers,

the threshold for the calculated correction in the grid points were set to $5\sigma=3m$. The results of automatic division into ‘Accepted area’ and ‘For further inspection’ are shown in Tab. 3.8.

Points accepted	3897 67%	
Outliers included	18 0.5%	
	Original DTM	Corrected DTM
Min. [m]	-4.10	-4.08
Max. [m]	10.89	9.82
Mean [m]	-0.05	-0.10
RMSE [m]	0.68	0.62
σ [m]	0.56	0.50

Tab. 3.8 The table shows how many percent of all 5778 investigated points appear in the category ‘Accepted’ (G) and how many outliers remain in that category when the correction is calculated as a mean of 25 surrounding points. At the same time thresholds for the calculated corrections ($corr < 3m$), correlation coefficient ($r > 0.3$), accuracy of LSM ($\sigma_{LSM} < 0.2$ pel), and standard deviation of an average corrected height ($\sigma_{mean} < 0.75m$) are applied.

The number of remaining outliers is relatively small. The average value was calculated without evaluating how many surrounding points per a grid point remain after applying thresholds for correlation coefficient and standard deviation of shift parameters of LSM. Therefore a new test was carried out.

C) Histogram of corrected heights of surrounding points

After applying the mentioned thresholds for the correlation coefficient and the standard deviation of shift parameters of LSM, a histogram of corrected heights of surrounding points was created and a peak of this histogram within a specified interval was sought. The interval was defined as the multiple of the expected standard deviation. Only in case that the interval contained at least a specified percentage of 25 points, the average height calculated from the heights within this interval was accepted. Different thresholds for minimal number of points within the interval, interval width and similarity measures were investigated. Some of the results with respect to the number of accepted points and outliers contained in those points are summarized in Tab. 3.9.

Interval		$\pm 3\sigma = \pm 2.25$ m		$\pm \sigma = \pm 0.75$ m		$\pm \sigma = \pm 0.75$ m	
T_r		0.3		0.3		0.5	
		Points accepted	Outliers	Points accepted	Outliers	Points accepted	Outliers
% of 25	60	73%	0.2%	48%	0.0%	32%	0.0%
	75	65%	0.1%	32%	0.0%	20%	0.0%
	90	47%	0.0%	12%	0.0%	7%	0.0%

Interval	$\pm 3\sigma = \pm 2.25$ m		$\pm 3\sigma = \pm 2.25$ m	
T_r	0.3		0.3	
σ_{LSM} [pel]	0.2		0.3	
	Points accepted	Outliers	Points accepted	Outliers
60 % of 25 points	29%	0.0%	50%	0.0%

Tab. 3.9 Results of the 'histogram' method. The tables show how many percent of all 5778 investigated points appear in the category 'Accepted' and how many outliers remain in that category when the correction is calculated by a 'histogram method'. The size of the interval is specified as well as the percentage of points that have to appear in this interval after applying thresholds for correlation coefficient T_r and accuracy of LSM.

As can be seen from Tab. 3.9, the number of outliers included in the area marked as 'acceptable' is minimal. Because the size of correction is not used as criterion for finding that area, the improvement of the DTM can be achieved (compare Tab. 3.10).

Thresholds				
interval	$\pm 3\sigma = \pm 2.25$ m			
% of 25 points	75		60	
T_r	0.3		0.3	
σ_{LSM} [pel]	0.3		0.3	
Points accepted	2141 37%		2892 50%	
Outliers included	1 0.0%		1 0.0%	
	O	C	O	C
σ [m] %h	0.71 (0.021)	0.48 (0.013)	0.78 (0.021)	0.49 (0.013)
Max. error [m] (absolute value)	8.28	2.74	13.33	2.74

O/C ...original/corrected DTM

Tab. 3.10 Improvement of the DTM in the 'accepted area' by the 'histogram' method

The results presented so far showed that the areas where the DTM is correct or improved could be found automatically. The number of remaining blunders and the size of area where the method works successfully differ depending on the thresholds of the applied criteria.

3.3.2.3 Results – Imagery 1:15 000

As mentioned at the beginning of the previous section, the problem with the systematic shift in the data set was overcome by a subtracting that shift from all original DTM heights. Its removal would be possible by a new orientation of the original images using correct control points. At the time when the tests were carried out, the images were about seven years old. Finding suitable natural control points in open land area would not be easy. A second set of new imagery at the scale of 1:15 000 was available. The possibilities of improvement the DTM model derived from 1:25 000 imagery by means of the larger scale and higher resolution images (1 pel \approx 15 μ m) was therefore investigated.

There are 2469 grid points in the two models 1:15 000 (see Fig. 3.5). Based on experience with imagery 1:25 000, following investigations were done:

- A) Calculation directly in the grid points and applying criteria for similarity measures
- B) Applying the histogram method
- C) Applying the histogram method with matching LR-RL (e.g. matching is done twice; first, the templates are derived from the left image for the first calculation and from the right image for the second calculation)

A) Calculation directly in the grid points and applying criteria for similarity measures

The calculation was done only for the northern model with 743 grid points. The comparison of this small sample with the reference data revealed values summarized in Tab. 3.11.

Number of points	743
Max. error [m]	9.62
Min. error [m]	-5.91
Mean [m]	1.64
RMSE [m]	1.82
σ [m]	0.78
Number of outliers	8% (60 points)

Tab. 3.11 Comparison of 743 heights of the DTM with reference values. The DTM will be checked by means of orthoimages derived from aerial images at the scale of 1:15 000.

For applying the tested correction method, the size of the templates was chosen $21 \times 21 \text{pel}^2$ and size of the search areas $23 \times 81 \text{pel}^2$ in order to be able to detect height changes up to 10 m that appeared in the model area (compare Tab. 3.11).

The corrected heights of grid points were derived and compared with the reference data directly without applying any thresholds for similarity measures or the size of correction. The results are summarized in Tab. 3.12.

	G_{ref}	Y_{ref}	R_{ref}
No. of points	623 (83.8%)	2 (0.3%)	118 (15.9%)

G_{ref}	Original DTM	Corrected DTM
Min. [m]	-5.19	-2.18
Max. [m]	8.40	2.22
Mean [m]	1.63	0.60
RMSE [m]	1.79	0.77
σ [m]	0.73	0.48

Tab. 3.12 Division of the 743 points of the corrected model DTM into categories G_{ref} , Y_{ref} and R_{ref} based on direct comparison with the reference data.

The reduction of systematic the shift and a higher accuracy of the corrected part of the model were achieved. In order to make the division of the model into the categories automatically, thresholds for similarity measures and the size of corrections were applied. The best results were obtained for thresholds presented in Tab. 3.13.

Points accepted	373 50%
Outliers included	2 0.5%
RMSE [m] %h	0.74 (0.020)
σ [m] %h	0.36 (0.010)
Max. error [m] (absolute value)	3.85

Tab. 3.13 The table shows how many percent of all 743 investigated points appear in the category 'Accepted' (G) and how many outliers remain in that category when applying thresholds for the calculated corrections ($\text{corr} < 2.25\text{m}$), correlation coefficient ($r > 0.5$) and accuracy of LSM ($\sigma_{\text{LSM}} < 0.2 \text{pel}$). RMSE and σ values are calculated from the corrected DTM heights.

Similarly to results presented in Tab. 3.6, matching in grid points and applying above mentioned thresholds the entire model was checked in about 50% of the area. In case of using images taken from a lower flying height and of a higher geometric resolution, the model was not only checked but also improved.

B) Histogram method

The height correction for each of 2469 grid points was derived from corrections calculated in 49 evenly distributed points in a square $4.05 \times 4.05 \text{ m}^2$ ($37 \times 37 \text{ pel}^2$) with its center in the grid point. The size of the templates of $15 \times 15 \text{ pel}^2$ and the size of the search areas of $17 \times 75 \text{ pel}^2$ was chosen. The criteria for correlation coefficient and accuracy of LSM were applied as well as criteria for histogram's interval grid and a number of points required within the interval. The overview of the results together with an example of DTM improvement is given in Tab. 3.14.

Thresholds								
interval	$\pm 3\sigma = \pm 2.25 \text{ m}$				$\pm \sigma = \pm 0.75 \text{ m}$			
% of 49 points	60		75		40		50	
T_r	0.3		0.3		0.3		0.3	
σ_{LSM} [pel]	0.3		0.3		0.3		0.3	
Points accepted	1096 44.4%		760 30.8%		1346 54.5%		1006 44.8%	
Outliers included	5 0.5%		0		1 0.0%		0	
	O	C	O	C	O	C	O	C
RMSE [m]	1.79	0.58	1.76	0.55	1.77	0.55	1.77	0.53
%h	(0.048)	(0.015)	(0.047)	(0.015)	(0.047)	(0.015)	(0.047)	(0.014)
Mean [m]	1.58	0.48	1.56	0.48	1.56	0.48	1.59	0.47
σ [m]	0.86	0.31	0.80	0.27	0.83	0.27	0.79	0.24
%h	(0.023)	(0.008)	(0.021)	(0.007)	(0.022)	(0.007)	(0.021)	(0.006)
Max. error [m] (absolute value)	5.91	3.23	4.89	1.93	5.91	3.01	5.91	1.76

O/C ...original/corrected DTM

Tab. 3.14 The table shows how many percent of all 2469 investigated points appear in the category 'Accepted' (G) and how many outliers remain in that category when the correction is calculated by the 'histogram method' from 49 surrounding points. At the same time thresholds for the width of histogram's interval, number of points required in the interval corresponding to the histogram's highest peak, correlation coefficient T_r and accuracy of LSM are applied.

An improvement of the model was achieved in about 50% of the area. Setting stronger limits of the applied thresholds brings a higher reliability but decreases the area where the tested method can be applied. The division of the points was done automatically.

C) Histogram method with matching LR-RL

The last test that was carried out is based on an additional geometric constrain. The height corrections are calculated twice. For the first time, the templates are derived from the left orthoimage and matched in the right orthoimage and for the second time vice versa. In the performed test no criteria for similarity measures were applied. There were only two requirements:

- at least 50% of neighboring points have to appear in the interval of histogram's highest peak
- the corrected height of a grid point derived from left-right and right-left method should not differ more than $\sigma\sqrt{2}=0.75\sqrt{2}m=1.06m$

The obtained results are presented in Tab. 3.15.

Points accepted	1639 66.4%	
Outliers included	0	
	O	C
RMSE [m] %h	1.76 (0.047)	0.56 (0.015)
Mean [m]	1.60	0.50
σ [m] %h	0.75 (0.020)	0.26 (0.007)
Max. error [m] (absolute value)	4.69	1.72

O/C ...original/corrected DTM

Tab. 3.15 The table shows how many percent of all 2469 investigated points appear in the category 'Accepted' and how many outliers remain in that category when applying histogram method and threshold for corrected height obtained by matching from left to right and vice versa. Any threshold for similarity measures was not used. The width of the interval was set to $\pm 3\sigma = \pm 2.25$ m. At least 50% of 49 points per grid point was required in the interval with the highest number of points. The difference of the corrected heights in a grid point derived from results of matching from left to right and vice versa should not differ more than $|dh_{LR}| = 1.06$ m. RMSE and σ values are calculated from the corrected DTM heights.

With respect to the amount of points where the DTM was checked and corrected it can be concluded that the last applied test was most successful. About 66% of points were accepted and standard deviation $\sigma = 0.26$ m (0.007%/h for the flying height $h=3750$ m) was achieved.

3.3.3 Evaluation of the method

The method of checking and correcting DTM based on determination of horizontal parallaxes between conjugate points in orthoimages brings improvement of the DTM. However, new outliers can be introduced. Therefore a direct application of the calculated corrections cannot be recommended. Most of the points where the method fails appear in areas of low texture, i.e. where conditions for area based matching are not fulfilled or where the grid point is surrounded by objects rising above the terrain. Exclusion of such areas is necessary. Other data sets such digital topographic maps, setting thresholds for similarity measures, applying geometric constraints can be used for that purpose.

A newly applied procedure based on finding corrections in the neighborhood of a grid point brought higher reliability to the decision where the DTM can be improved and where the method fails. The influence of single objects rising above the terrain, for example single trees or bushes can be decreased in this way. The suggested procedure worked well within the test data, i.e. open land and relatively flat terrain. The distance between the surrounding points and the grid point could be several meters. In hilly or mountainous terrain the distribution of surrounding points around a grid point must be watched when using such a procedure.

Using imagery of the same scale as it was used for DTM derivation brought an improvement of the DTM about 30% (see Tab. 3.10). In order to discover and eliminate systematic errors, new and accurate orientation of the stereopair must be carried out. Using images of a larger scale and better geometric resolution (a smaller pixel size) resulted in a DTM improvement of 65% (see Tab. 3.15). On the other hand using such an approach requires more resources and the question is whether it is not more economic to derive a new DTM automatically.

Finding conjugate points directly in original or normalized images would be an alternative to the studied method. When using orthoimages, the most time consuming part is matching itself. Opposite to using original images, all transformations between image and object coordinate systems are avoided. In case that the flight was done parallel to one axis of the reference co-ordinate system, searching along epipolar lines is equal to searching along rows or columns of the orthoimage.

The results showed that the automatic check and possibly also correction of the existing DTM by a proposed method is successful. The DTM is divided into two parts. The first one, that does not contain outliers at all or contains only a small percentage of them (max. 0.5%, see e.g. Tab. 3.10). The second part must be checked (and improved) by another method. If the second part is checked manually, an operator can be automatically navigated to the points that have to be measured. The results of the DTM check by the studied method can also be presented in the form of a plot as it is shown in Fig. 3.13. Most of the points where the method failed are in areas with a very low texture. There is also a possibility of correcting a single red point surrounded by green points by filtering. All the applied thresholds were found empirically and are valid for the test data set. Further investigations have to be carried out in order to find thresholds for other types of the landscape.

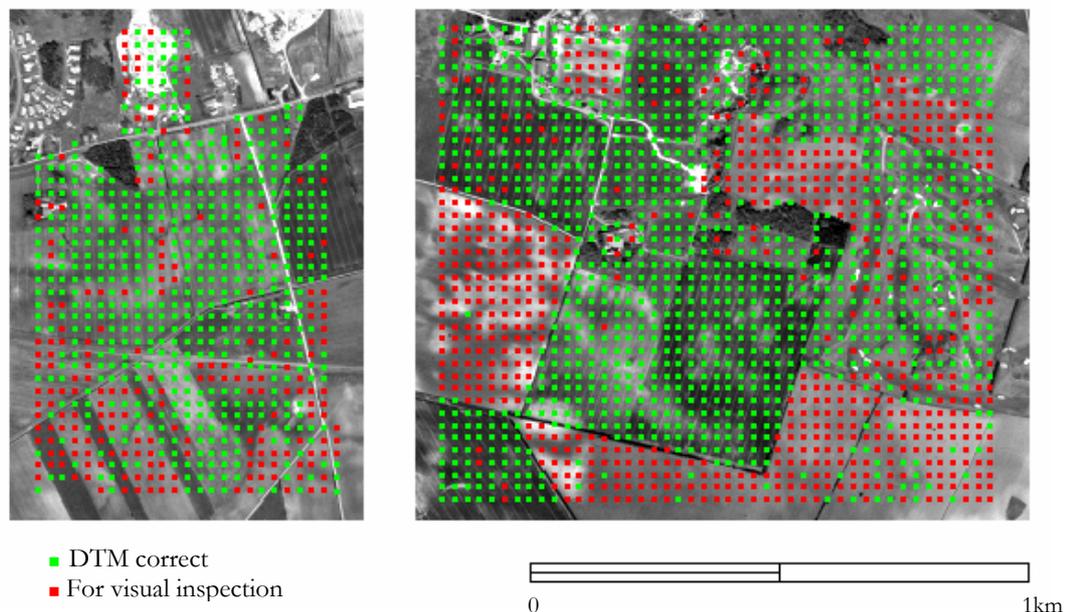


Fig. 3.13 Overview of the checked DTM by means of orthoimages derived from 1:15 000 imagery. Green dots show points that were corrected (about 66%), red dots correspond to points where another checking method must be used.

Conclusion

Automatic measurement in digital images is the basic procedure in modern photogrammetry. It has been incorporated in many applications such as image orientation, DTM generation, refinement and update of vector data etc. In comparison to the manual approach, it gives a possibility to obtain higher accuracy in shorter time. Due to a relatively high amount of incorrect measurements, both of these advantages can only be achieved only if the systems providing tools for an automatic measurement are able to deal with these erroneous data.

The presented work summarises procedures for automatic finding conjugate points in overlapping images. Most attention is paid to area based methods. The application of these methods for orientation of an aerial image, quality control of an orthoimage and a digital terrain model are rather new although some investigations have been carried out during the last ten years.

The main achievements of the thesis can be summarised as follows:

1. The relation between similarity measures namely the correlation coefficient, image distance, and mutual information is described. The use of all three measures for area based matching is evaluated. The results of experiments lead to the following conclusion. In case of minimal radiometric differences between image patches there are no differences in the position of the best fit obtained by means of the correlation coefficient, image distance and mutual information. In case of aerial images or orthoimages when the intensity values in the image patches vary due to different illumination, viewing angle or temporal changes, it is only the correlation coefficient that does not require any further pre-processing in order to minimise the amount of mismatches. All three measures are scale and rotation dependent.
2. Thresholds for different similarity measures and their combination for reducing outliers in area based matching are presented. It is shown that such thresholds are not sufficient for the elimination of all outliers and that additional geometric constrains and robust adjustment must be applied for their considerable reduction or complete removal.
3. Improvements of the method of automatic orientation of aerial images based on existing data sets, namely an orthoimage, topographic map and DTM, is achieved especially in the accuracy and the automation of all processes from extraction of control information to computing orientation parameters. A hierarchical approach (an image and object pyramid)

is applied in order to improve the approximations of orientation parameters. The accuracy of the automatically derived orientation parameters corresponds to a co-ordinate error of 23 μm (0.9 pel) in the image and 0.05‰ of the flying height. A new orthoimage is derived using obtained orientation parameters and an existing DTM.

4. The quality control of the new orthoimage is carried out automatically by comparison with the existing orthoimage. The comparison with the map is based on manual measurements but possibilities of an automation of this process are mentioned as well. Error propagation in the studied orthoimage production chain is described and the obtained differences are explained. The importance of good correspondence of input data is emphasised. The agreement between the new and the existing orthoimage of $\sigma_{XY}=0.3$ pel can be achieved.
5. The method of DTM checking and correcting by means of overlapping orthoimages is analysed and developed further. The disadvantages of the method such as its failure in areas of low texture are pointed out. Thresholding of several similarity measures and additional geometric constrains are necessary to apply in order to exclude outliers. The new approaches based on the calculations of height corrections in the neighbourhood of the grid points are investigated. The result of the developed approach is an automatic division of all DTM points into two categories. In the first category there are points where geometric constrains and requirements for similarity measures are fulfilled. The elevations of these points do not need any further improvements. If outliers appear in this category of points, their amount does not exceed the number acceptable for normal error distribution. The points in the second category must be checked by means of other methods. In case of DTM automatically derived from aerial images, the best results are achieved when the DTM is checked by means of the imagery of larger scale and higher geometric resolution than it was created.
6. The use of a topographic database for excluding areas that are not suitable for correlation (forests, dense urban areas) is shown.
7. Suggestions for a further improvement of the studied methods are given. The method of automatic orientation of images should be tested on data sets with different scale (both satellite imagery and large scale imagery) and for a block of images (combining automatic measurement of both tie points and control points). Other time invariant objects as houses or roads should be considered as control information. The DTM correction method should also be tested on different landscapes and not only with DTMs derived automatically, e.g. with a relatively small amount of outliers, but also with DTMs derived from contour lines or laser scanning.

The main contribution of this thesis is in bringing the method of automatic orientation of images including the comparison of a new orthoimage with the existing data to the stage of full functionality. It needs some professional programming to bring it to efficient production. Regarding the method of DTM checking and correcting, the presented work is the first step for fulfilling the requirements of the mapping agencies for a method which checks most of the DTM area and leads the operator to the problem areas.

All the computations concerning image matching, DTM correction and spatial resection were carried out by means of developed experimental software. The developed programs that can be found in the attached CD were used not only for the presented research work but also for education purposes. Moreover, experience with the discussed topics was used for development of learning programs AutoOrient and LDIPInter2.

References

- Ackermann, F., Schneider, W., 1992, Experience with automatic DEM generation, International Archives for Photogrammetry and Remote Sensing, 30(B4), pp. 986-989
- Ackermann, F., 1996a, Some considerations about automatic digital aerial triangulation, Photogrammetric Week '96, [<http://phot.epfl.ch/workshop/wks96/tocwks96.html>]
- Ackermann, F., 1996b, Techniques and strategies for DEM generation, In: Digital Photogrammetry: An Addendum to the Manual of Photogrammetry, American Society for Photogrammetry and Remote Sensing, ISBN 1-57083-037-1
- Albertz, J., Kreiling, W., 1975, Photogrammetric guide, 2nd edition, Herbert Wichmann Verlag, Karlsruhe, ISBN 3-87907-026-1
- Atkinson, K. B., 1996, Close range photogrammetry and machine vision, Whittles Publishing, Caithness, UK, ISBN 1-870325-46-X
- Brandt-Lavridsen, O., 1993, Fotogrammetri, 9th edition, Lab. for fotogrammetri og landmåling, Aalborg Universitetscenter
- Calitz, M. F., Rüther, H., 1994, L-norm methods in area based matching, International Archives for Photogrammetry and Remote Sensing, 30 (5), pp. 22-27
- Fischer, M. A., Bolles, R. C., 1981, Random sample consensus: A paradigm for model fitting with application to image analysis and automated cartography, Graphics and Image processing, Vol. 24, No. 6, pp. 381-395
- Förstner, W., 1996, A feature based algorithm for image matching, International Archives for Photogrammetry and Remote Sensing, 26(3/2), pp.150-166
- Gao, Y., Krakiwsky, E. J., Czompo, J., 1992, Robust testing procedure for detection of multiple blunders, Journal of Surveying Engineering, Vol. 118, No.1, pp. 11-23
- Gooch, M. J., Chandler, J. H., 2000, A novel error technique for automatically generated digital elevation models, International Archives for Photogrammetry and Remote Sensing, 33(B4/1), pp. 302-309
- Gülch, E., 1995, Automatic control point measurement, Photogrammetric Week '95, Fritsch, D., Hobbie, D. (Eds.), pp. 185-196, Wichmann Verlag, Heidelberg
- Habib, A., Kelley, D., 2001, Automatic relative orientation of large scale imagery over urban areas using Modified Integrated Hough Transform, ISPRS Journal of Photogrammetry & Remote Sensing 56, pp. 29-41
- Hahn, M., 1997, Automatic control point measurement, Photogrammetric Week '97, Fritsch, D., Hobbie, D. (Eds.), pp. 115-126, Wichmann Verlag, Heidelberg

- Heipke, C., 1996, Overview of image matching techniques, Photogrammetric Week '96, [<http://phot.epfl.ch/workshop/wks96/tocwks96.html>]
- Heipke, C., 1997, Automation of interior, relative, and absolute orientation, ISPRS Journal of Photogrammetry & Remote Sensing 52, pp. 1-19
- Heipke, C., Jakobsen, K., Wegmann, E., 2002, Integrated sensor orientation – Test report and workshop proceedings, OEEPE, Official publication No. 43
- Heipke, C., Pakzad, K., Willrich, F., 2004, Integration of geodata and imagery for automated refinement and update of spatial databases, ISPRS Journal of Photogrammetry & Remote Sensing 58, pp. 127-128
- Helava, U. V., 1978, Digital correlation in photogrammetric instruments, Photogrammetria, Vol. 34, No. 1, pp. 19-41
- Helava, U. V., 1988, Object-space least-squares correlation, Photogrammetric Engineering and Remote Sensing, Vol. 54, No. 6, pp. 711-714
- Höhle, J., 1999a, Automatic orientation of the aerial images on database information, OEEPE, Official publication Nr. 36, pp. 71-117
- Höhle, J., 1999b, Orientation of aerial images by means of existing orthoimages and height models – results from experiments with OEEPE test material, OEEPE, Official publication Nr. 36, pp. 159-165
- Höhle, J., 2001, Automated georeferencing of aerial images, Proceedings of the 8th Scandinavian research conference ScanGIS'2001, pp. 115-123
- Höhle, J., Potůčková, M., 2001, Steps to automated orthoimage production, Proceedings of International Symposium 'Geodetic, Photogrammetric and Satellite technologies – Development and Integrated Application', Sofia
- Höhle, J., 2003, Automatic orientation of aerial images on topographic databases – an overview, in Automatic Georeferencing of Aerial Images by Means of Topographic Database Information, Department of Development and planning, Aalborg University 2003
- Honkavaara, E., Kaartinen, H., Kuittinen, R. Huttunen, A., Jaakkola, J., 1999, Quality of FLPIS orthophotos, Reports of the Finnish Geodetic Institute, 99:1, Helsinki
- Huber, P. J., 1981, Robust statistics, John Wiley & Sons, Inc., ISBN 0-471-41805-6
- Jacobi, O., 1994, Digitale højdemodeller, In: GIS i Danmark, Balstrøm, T., Jacobi, O., Sørensen, E. B. (Eds.), pp. 61-71, Teknisk Forlag A/S, København, ISBN 87-571-1775-6
- Jähne, B., 1997, Practical handbook on image processing for scientific applications, CRC Press LLC, ISBN 0-8493-8906-2

- Jedryczka, R., 1999, Semi-automatic exterior orientation using vector map data, OEEPE, Official publication Nr. 36, pp. 133-137
- Jørgensen, P. C., Frederiksen, P., Kubik, K., Weng, W., 1984, Ah, robust estimation!, International Archives for Photogrammetry and Remote Sensing, 25(A3a), pp. 269-277
- Juhl, J., 1984, The 'Danish method' of weight reduction for gross error detection, International Archives for Photogrammetry and Remote Sensing, 25(A3a), pp. 468-472
- Karras, G. E., Mavogenneas, N., Mavrommati, D., Tsikonis, N., 1998, Test on automatic DEM generation in a digital photogrammetric workstation, International Archives for Photogrammetry and Remote Sensing, 32 (2), pp. 136-139
- Kasser, M., Egels, Y., 2002, Digital photogrammetry, Taylor & Francis, London, ISBN 0-748-40945-9
- Karjalainen, M., Kuittinen, R., 1999, Interactive exterior orientation using linear features from vector map, OEEPE, Official publication Nr. 36, pp. 127-131
- Krärup, T., Juhl, J., Kubik, K., 1980, Götterdämmerung over least squares adjustment, International Archives for Photogrammetry and Remote Sensing, 23(B3), pp. 369-377
- Kraus, K., 2000, Photogrammetry, Vol.1., 4th edition, Verlag H. Stam GmbH, Köln, ISBN 3-427-78684-6
- Kraus, K., 1997, Photogrammetry, Vol.2., 4th edition, Fer. Dümmlers Verlag, Bonn, ISBN 3-427-78694-3
- Krupnik, A., 1998, Automatic detection of erroneous areas in automatic surface reconstruction, International Archives for Photogrammetry and Remote Sensing, 32 (3/1), pp. 132-137
- Kubik, K., Weng, W., Frederiksen, P., 1984, Oh, gross errors!, International Archives for Photogrammetry and Remote Sensing, 25(A3a), pp. 279-289
- Läbe, T., Ellenbeck, K.H., 1996, 3D-wireframe models as ground control points for the automatic exterior orientation, International Archives for Photogrammetry and Remote Sensing, 31(B2), pp. 218-223
- Läbe, T., 1999, Automatic orientation of aerial images, Task A – Experiences with AMOR, OEEPE, Official publication No. 36, pp. 119-125
- Läbe, T., 2003, Using automatic orientation in an orthophoto production process, In: Automatic georeferencing of aerial images by means of topographic database information, Höhle, J. (Ed.), Department of development and planning, Aalborg University

- Lacey, A. J., Pinitkarn, N., Thacker, N. A., An evaluation of the performance of RANSAC algorithm for stereo camera calibration, Proceedings to BMVC 2000, Bristol [<http://www.bmva.ac.uk/bmvc/2000/contents.htm>]
- Larsen, J. N., 1998, Improving DEM by integrating criteria, International Archives for Photogrammetry and Remote Sensing, 32 (2), pp. 177-183
- Li, D., Wang, S., Li, R., 1996, Automatic quality diagnostics in DTM generation by digital image matching techniques, Geomatica, Vol. 50, No. 1, pp. 65-73
- Lühmann, T., Altrogge, G., 1986, Interest-operator for image matching, International Archives for Photogrammetry and Remote Sensing, 26(3/2), pp.459-474
- Lühmann, T., 2000, Nahbereichsphotogrammetrie: Grundlagen, Methoden und Anwendungen, Herbert Wichmann Verlag, Heidelberg, ISBN 3-87907-321-X
- Maes, F., Collignon, A., Vandermeulen, D., Marchal, G., Suetens, P., 1997, Multi-modality image registration by maximization of mutual information, IEEE Transactions on medical imaging, No. 16, pp. 187-198
- Mikhail, E. M., Garcie, G., 1981, Analyses and adjustment of survey measurements, Van Nostrand Reinhold Company, Inc., New York, ISBN 0-442-25369-9
- Mikhail, E. M., Bethel, J. S., McGlone J. C., 2001, Introduction to modern photogrammetry, John Wiley & Sons, Inc., ISBN 0-471-30924-9
- Norvelle, F. R., 1996, Using iterative orthophoto refinements to generate and correct digital elevation models (DEM's), In: Digital Photogrammetry: An Addendum to the Manual of Photogrammetry, American Society for Photogrammetry and Remote Sensing, ISBN 1-57083-037-1
- Paszotta, Z., 1999, Matching orthoimages and direct method determining exterior orientation elements, OEEPE, Official publication Nr. 36, pp. 145-150
- Paszota, Z., 2000, Method of exterior orientation of aerial images by matching orthoimages, Wydawnictwo Uniwersytetu Warmińsko-mazurskiego, Olsztyn
- Pedersen, B. M., 1996, Automated measurement of ground control objects in large scale aerial photographs, International Archives for Photogrammetry and Remote Sensing, 31(B3), pp. 633-637
- Pedersen, B. M., 1999a, OEEPE test on automatic orientation – a solution from Aalborg, OEEPE, Official publication Nr. 36, pp. 139-144
- Pedersen, B. M., 1999b, Automatisk måling i digitale fotogrammetriske billeder, Landinspektøren, 1/99

- Potůčková, M., 2003, Experience with automatic orientation from different data sets, in Automatic Georeferencing of Aerial Images by Means of Topographic Database Information, Department of Development and planning, Aalborg University 2003
- Rosenholm, D., 1986, Accuracy improvement in digital matching, Photogrammetric reports No. 52, The Royal Institute of Technology, Department of Photogrammetry, Stockholm, Sweden
- Rousseeuw, P. J., Leroy, A. M., 1987, Robust regression and outlier detection, John Wiley & Sons, Inc., ISBN 0-471-85233-3
- Russ, J. C., 1999, The image processing handbook, 3rd ed., CRC Press LLC, ISBN 3-540-64747-3
- Schenk, T., Li, J. C., Toth, C., 1991, Towards an autonomous system for orienting digital stereopairs, Photogrammetric Engineering and Remote Sensing, Vol. 57, No. 8, pp. 1057-1064
- Schenk, T., Hoffmann, O., 1986 Stereomatching using line segments of zero crossings, International Archives for Photogrammetry and Remote Sensing, 26(3/2), pp.602-607
- Schenk, T., 1999, Digital photogrammetry, TerraScience
- Schickler, W., 1992, Feature matching for outer orientation of single images using 3D wireframe controlpoints, International Archives for Photogrammetry and Remote Sensing, 30(B3/III), pp.591-598
- Schwarz, C. R., Kok, J., 1993, Blunder detection and data snooping in LS and robust adjustment, Journal of Surveying Engineering, Vol. 119, pp. 127-136
- Shahrabi, B. A., 2000, Automatic recognition and 3D reconstruction of buildings from digital imagery, Deutsche Geodätische Kommission, München
- Shan, J., 1999, Automatic exterior orientation with orthoimage and DTM, OEEPE, Official publication Nr. 36, pp. 151-157
- Tang, L., Heipke, Ch., 1996, Automatic relative orientation of aerial images, Photogrammetric Engineering and Remote Sensing, Vol. 62, No. 1, pp. 47-55
- TOP10DK specifications, 2001, The National Survey and Cadastre, Copenhagen
- Wehr, A., Lohr, U., 1999, Airborne laser scanning – introduction and overview, ISPRS Journal of Photogrammetry & Remote Sensing 54, pp. 68-82
- Werner, H., 1984, Automatic gross error detection by robust estimators, International Archives for Photogrammetry and Remote Sensing, 25(A3a), pp. 1101-1108
- Wind, M., 1998, Investigation into the automatic generation of heights of different Danish landscape types, GIS-Between Visions and Applications, ISPRS Commission Symposium, Vol. 32 Part 4, pp. 662 - 668

- Wind, M., 2001, Accuracy of automatically collected height data over Denmark, the 20th International Cartographic Conference, Beijing China
- Yu, Z., Jiang, P., 2001, Distance, correlation and mutual information among portraits of organisms based on complete genomes, Physics letters A 286, pp. 34-46

Learning programs developed by Prof. J. Höhle, M. Potůčková, and E. Jensen:

[<http://ldipinter2.plan.aau.dk>]

[<http://www.internetsurvey.dk/lp>]

Acknowledgement

I would like to express my sincere gratitude to several people whose support and guidance made this thesis come true.

I would like to thank Professor Bohuslav Veverka for his kind supervision during my Ph.D. studies, for giving me freedom to realise all my plans and ideas and for helping and encouraging me with discovering the world.

I wish to thank to all my colleagues at the Laboratory for Geoinformatics, Aalborg University for their help, patience, and understanding during my work on this project. My special thanks go to Professor Joachim Höhle for his great support from the very first moment I started to work at Aalborg University. I would like to thank him for introducing me to the field of photogrammetry and image processing, for his guidance and help with my research work, for all discussions, comments and advice.

I would like to acknowledge the National Survey and Cadastre in Copenhagen that provided data sets for testing the method of automatic orientation of aerial images.

Last but not least, I want to thank my parents for the encouragement to finish this work, for their great support, care and patience during all my studies.

Appendix A

Spatial resection with robust adjustment

Spatial resection with robust adjustment

In this appendix, spatial resection solved by the ‘Danish’ robust adjustment method is described. The same solution is applied in the developed MATLAB[®] function ***r_robust.m***. The function was used for the calculation of orientation parameters of an aerial image in the tests carried out in chapter 2. It can be found in together with other functions mentioned in following paragraphs in Appendix C.

The meaning of spatial resection is to determine the orientation parameters, i.e. position of the perspective centre X_0, Y_0, Z_0 and rotations ω, φ , and κ of a single image when positions of at least three ground control points (GCPs) are known both in the image and the object coordinate systems. The calculation of resection is based on collinearity equations A.1 (see also Fig. 1.3 in the chapter 1.1).

$$\begin{aligned} x' &= x'_0 - c \frac{r_{11}(X - X_0) + r_{21}(Y - Y_0) + r_{31}(Z - Z_0)}{r_{13}(X - X_0) + r_{23}(Y - Y_0) + r_{33}(Z - Z_0)} = x'_0 - c \frac{U}{W} \quad \text{or} \quad F_1 = x' - x'_0 + c \frac{U}{W} = 0 \\ y' &= y'_0 - c \frac{r_{12}(X - X_0) + r_{22}(Y - Y_0) + r_{32}(Z - Z_0)}{r_{13}(X - X_0) + r_{23}(Y - Y_0) + r_{33}(Z - Z_0)} = y'_0 - c \frac{V}{W} \quad \text{or} \quad F_2 = y' - y'_0 + c \frac{V}{W} = 0 \end{aligned} \tag{A.1}$$

x', y' image co-ordinates of a GCP

X, Y, Z object co-ordinates of a GCP

X_0, Y_0, Z_0 object co-ordinates of the perspective centre

r_{ij} elements of the rotation matrix containing angles ω, φ , and κ (Kraus, 2000)

c principle distance

x'_0, y'_0 image co-ordinates of the principal point

Before calculating orientation parameters, the image co-ordinates are corrected for lens distortion, earth curvature, and atmospheric refraction. The values of averaged radial lens distortion errors from a camera calibration report are input values of the function ***rad_dist.m***. The corrections at each observed point are derived by means of linear interpolation between calibrated values. Tangential lens distortion is neglected. The function ***earth_ref.m*** corrects image co-ordinates for earth curvature and atmospheric refraction by means of formula A.2 (Kraus, 2000, Brande-Lavridsen, 1993). The correction for earth curvature is rather simple and made on an assumption of relatively small height differences in an area covered by an image. It is also possible to calculate the correction directly in the object space when transforming ground co-ordinates from the reference system to the tangential system (Kraus, 1997).

$$dr_{ea} = dr_e - dr_a \quad (A.2)$$

Correction for earth curvature dr_e

Correction for atmospheric refraction dr_a

$$dr_e = \frac{Hr'^3}{2Rc^2}$$

$$dr_a = r' \left(1 + \left(\frac{r'}{c} \right)^2 \right) K$$

r' radial distance of the measured point to the principal point of the image

c principle distance

H flying height

R Earth radius

K refraction coefficient (Kraus 2000)

If more than three GCPs are given, least squares adjustment can be applied. The calculation requires linearization of collinearity equations and approximations of unknown parameters $\mathbf{x}^0 = [X_0^0, Y_0^0, Z_0^0, \omega^0, \varphi^0, \kappa^0]$. It is carried out as an iterative process where the calculated corrections $\Delta = [dX, dY, dZ, d\omega, d\varphi, d\kappa]$ are added to orientation parameters used in the previous step:

$$\begin{aligned} X_0^{i+1} &= X_0^i + dX^{i+1} & \omega_0^{i+1} &= \omega_0^i + d\omega^{i+1} \\ Y_0^{i+1} &= Y_0^i + dY^{i+1} & \varphi_0^{i+1} &= \varphi_0^i + d\varphi^{i+1} & i = 0, \dots, k \\ Z_0^{i+1} &= Z_0^i + dZ^{i+1} & \kappa_0^{i+1} &= \kappa_0^i + d\kappa^{i+1} \end{aligned}$$

The calculation stops as soon as the corrections of orientation parameters are smaller than given thresholds. The number of iterations is limited for case of slow convergence or divergence of the solution.

The algorithms mentioned in (Mikhail, 2001) and (Albertz, 1975) use only image co-ordinates of GCPs as observations. Their object co-ordinates are considered known and error-free as well as the principle distance c and image co-ordinates of the principal point x'_0, y'_0 . In the function **r_robust.m**, object co-ordinates of GCPs are also included as observations. The linearized observation equations can be then written as

$$\begin{bmatrix} \frac{\partial F_1}{\partial \mathbf{l}} \\ \frac{\partial F_1}{\partial \mathbf{x}} \\ \frac{\partial F_2}{\partial \mathbf{l}} \\ \frac{\partial F_2}{\partial \mathbf{x}} \end{bmatrix} \mathbf{v} + \begin{bmatrix} \frac{\partial F_1}{\partial \mathbf{x}} \\ \frac{\partial F_1}{\partial \mathbf{l}} \\ \frac{\partial F_2}{\partial \mathbf{x}} \\ \frac{\partial F_2}{\partial \mathbf{l}} \end{bmatrix} \Delta = \begin{bmatrix} -F_1(\mathbf{l}^0, \mathbf{x}^0) \\ -F_2(\mathbf{l}^0, \mathbf{x}^0) \end{bmatrix}$$

$$\mathbf{A}\mathbf{v} + \mathbf{B}\Delta = \mathbf{f} \quad (\text{A.3})$$

n number of GCPs

\mathbf{l} vector of observations
_{5n}

\mathbf{x} vector of unknown parameters
_{6,1}

\mathbf{v} vector of image and object co-ordinate residuals $\mathbf{v} = [\mathbf{v}_x, \mathbf{v}_y, \mathbf{v}_x, \mathbf{v}_y, \mathbf{v}_z]$
_{5n,1} _{5n,1}

\mathbf{A} matrix of partial derivatives of functions F with respect to each observed co-ordinate
_{2n,5n}

\mathbf{B} matrix of partial derivatives of functions F with respect to each orientation parameter
_{2n,6}

Δ vector of corrections to approximations of orientation parameters
_{6,1}

\mathbf{f} vector of constant terms (differences between measured image co-ordinates and
_{2n,1} values calculated from collinearity equations using approximations of orientation parameters)

The partial derivatives in matrices \mathbf{A} and \mathbf{B} can be found in (Kraus, 2000).

The condition of least squares adjustment $\mathbf{v}^T \mathbf{P} \mathbf{v} = \min$ leads to normal equations A.4

$$\mathbf{B}^T (\mathbf{A}^T \mathbf{P} \mathbf{A})^{-1} \mathbf{B} \Delta = \mathbf{B}^T (\mathbf{A}^T \mathbf{P} \mathbf{A})^{-1} \mathbf{f} \quad (\text{A.4})$$

$$\mathbf{B}^T \mathbf{W} \mathbf{B} \Delta = \mathbf{B}^T \mathbf{W} \mathbf{f}$$

with the solution

$$\Delta = (\mathbf{B}^T \mathbf{W} \mathbf{B})^{-1} \mathbf{B}^T \mathbf{W} \mathbf{f}$$

\mathbf{P} is a weight matrix

$$\mathbf{P} = \text{diag} \left(\frac{\sigma_0^2}{\sigma_{x'1}^2}, \dots, \frac{\sigma_0^2}{\sigma_{x'n}^2}, \frac{\sigma_0^2}{\sigma_{y'1}^2}, \dots, \frac{\sigma_0^2}{\sigma_{y'n}^2}, \frac{\sigma_0^2}{\sigma_{x1}^2}, \dots, \frac{\sigma_0^2}{\sigma_{xn}^2}, \frac{\sigma_0^2}{\sigma_{y1}^2}, \dots, \frac{\sigma_0^2}{\sigma_{yn}^2}, \frac{\sigma_0^2}{\sigma_{z1}^2}, \dots, \frac{\sigma_0^2}{\sigma_{zn}^2} \right)$$

If the value of an a priori standard deviation of unit weight σ_0 is set equal to a standard deviation of the image co-ordinates, i.e. $\sigma_0 = \sigma_{x'} = \sigma_y$, then

$$\mathbf{P} = \text{diag} (1, \dots, 1, 1, \dots, 1, \frac{\sigma_0^2}{\sigma_{x1}^2}, \dots, \frac{\sigma_0^2}{\sigma_{xn}^2}, \frac{\sigma_0^2}{\sigma_{y1}^2}, \dots, \frac{\sigma_0^2}{\sigma_{yn}^2}, \frac{\sigma_0^2}{\sigma_{z1}^2}, \dots, \frac{\sigma_0^2}{\sigma_{zn}^2})$$

σ_x , σ_y , and σ_z are the a priori standard deviations of measured object co-ordinates.

In order to decrease influence of possible outliers, the weight matrix \mathbf{P} of measured co-ordinates changes with each iteration. The original weight P_{ii} ($i=1, \dots, 5n$) is multiplied with a factor pr (Juhl, 1984):

1 st iteration	pr=1
2 nd and 3 rd iteration	pr=exp(-0.05(v _i /σ _i) ^{4.4})
following iterations	pr= exp(-0.05(0.6v _i /σ _i) ^{6.0})

σ_i are a priori standard deviations of measured co-ordinates. After each iteration step, the vector of residuals in observed co-ordinates **v** are calculated together with a posteriori standard deviation of a unit weight S₀.

$$\begin{aligned}\mathbf{v} &= \mathbf{P}^{-1} \mathbf{A}^T \mathbf{W} (\mathbf{B} \Delta - \mathbf{f}) \\ S_0^2 &= \mathbf{v}^T \mathbf{P} \mathbf{v} / (2n-6)\end{aligned}\tag{A.5}$$

The whole calculation stops as soon as corrections of orientation parameters fulfil convergence criteria or the specified number of iterations is exceeded (k=20).

Standard deviations of the orientation parameters **S**_Δ and root mean square errors of co-ordinate residuals are calculated after the last iteration as a final evaluation of the whole calculation. The outliers, i.e. points with a weight p→0 are excluded from the calculation of the root mean square errors in co-ordinates X, Y, and Z.

$$\begin{aligned}\mathbf{S}_\Delta^2 &= S_0^2 (\mathbf{B}^T \mathbf{W} \mathbf{B})^{-1} \\ \text{RMSE}_X &= (\mathbf{v}_X^T \mathbf{v}_X / n_p) \quad \text{RMSE}_Y = (\mathbf{v}_Y^T \mathbf{v}_Y / n_p) \quad \text{RMSE}_Z = (\mathbf{v}_Z^T \mathbf{v}_Z / n_p)\end{aligned}$$

v_X, **v**_Y, **v**_Z ... vectors of residuals in object co-ordinates X, Y, Z in n_p ground control points
n_p = n - n_{out} ... n_{out} is the number of outliers discovered by the adjustment procedure

Appendix B

B.1 Correlation function

B.2 Least squares matching

B.1 Correlation function

A position of the best fit between a template and a search area calculated by means of correlation coefficient is determined as an integer value, i.e. even in a case of an errorless match its accuracy is not better than ± 0.5 pel. In order to obtain subpixel accuracy, a continuous function that fits the correlation coefficient values around the position of the best fit is established and the position of the maximum of such a function is considered as an improved position of the best fit. Theoretically, any continuous function reaching only one maximum around the highest correlation coefficient value can be chosen. A normal distribution curve is one of possibilities. However, a discrete correlation function is approximated in a relatively small area and it gives from a computational point of view a reason for a simpler function, namely a 2nd order polynomial. Its coefficients are found by least squares adjustment with an observation equation (B.1) – adapted from (Kraus, 1997):

$$\bar{r}_{coef.} = r_{coef.} + \nu = a_0 + a_1 r + a_2 c + a_3 r c + a_4 r^2 + a_5 c^2 \quad (B.1)$$

$r_{coef.}$ calculated correlation coefficient

$\bar{r}_{coef.}$ adjusted correlation coefficient

r, c pixel co-ordinates

a_i coefficients of a 2nd order polynomial

ν residual

In matrix notation:

$$\mathbf{v} = \mathbf{A} \mathbf{dx} - \mathbf{l} \quad (B.2)$$

\mathbf{dx} vector of unknowns coefficients

\mathbf{A} design matrix containing pixel co-ordinates and their multiplication

\mathbf{l} vector of calculated correlation coefficients

The condition of least squares leads to a solution (B.3)

$$\mathbf{dx} = (\mathbf{A}^T \mathbf{P} \mathbf{A})^{-1} \mathbf{A}^T \mathbf{P} \mathbf{l} \quad (B.3)$$

\mathbf{P} is a weight matrix. For further calculations it assumed that $\mathbf{P} = \mathbf{I}$ (identity matrix).

After the coefficients of the polynomial are known, the position of its maximum is calculated by solving equation (B.4) obtained by differentiating the equation (B.1):

$$\begin{pmatrix} \partial r_{coef.} / \partial r \\ \partial r_{coef.} / \partial c \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + \begin{pmatrix} 2a_4 & a_3 \\ a_3 & 2a_5 \end{pmatrix} \begin{pmatrix} r_{max} \\ c_{max} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (B.4)$$

$$r_{max} = \frac{-2a_1 a_5 + a_2 a_3}{4a_4 a_5 - a_3^2} \quad c_{max} = \frac{-2a_2 a_4 + a_1 a_3}{4a_4 a_5 - a_3^2}$$

The variances of derived coefficients of the polynomial are diagonal elements of a covariance matrix Σ_x (equation B.5).

$$\Sigma_x = \sigma_0^2 \mathbf{N}^{-1} = \sigma_0^2 (\mathbf{A}^T \mathbf{P} \mathbf{A})^{-1} \quad (B.5)$$

$\sigma_0^2 = \mathbf{v}^T \mathbf{P} \mathbf{v} / (m-n)$ is an a posteriori variance. The number of observations m must fulfil a condition $m > n$, where the number of unknowns $n=6$.

The standard deviations σ_r and σ_c of the determined position of the best fit are derived by a law of error propagation (formula B.6):

$$\sigma_r^2 = \mathbf{B}_r \Sigma_x \mathbf{B}_r^T \quad (B.6a)$$

$$\sigma_c^2 = \mathbf{B}_c \Sigma_x \mathbf{B}_c^T \quad (B.6b)$$

$$\mathbf{B}_r = [\partial r_{\max} / \partial a_0, \partial r_{\max} / \partial a_1, \dots, \partial r_{\max} / \partial a_5]$$

$$\mathbf{B}_c = [\partial c_{\max} / \partial a_0, \partial c_{\max} / \partial a_1, \dots, \partial c_{\max} / \partial a_5]$$

In order to simplify the calculation, a local pixel co-ordinate system is established with its origo in the centre of a pixel with the highest correlation coefficient. A solution with nine observations is shown in Fig. B.1.

Correlation coefficient values and its maxima

- within search area

		col		
		31	32	33
row	29	0.61	0.72	0.68
	30	0.67	0.79	0.74
	31	0.61	0.73	0.69

- local pixel co-ordinate system

		col		
		-1	0	1
row	-1	0.61	0.72	0.68
	0	0.67	0.79	0.74
	1	0.61	0.73	0.69

Observation equations

\mathbf{v}	=	\mathbf{A}	\mathbf{dx}	-	$\mathbf{1}$				
v ₋₁₋₁		1	-1	-1	1	1	1		
v ₋₁₋₀		1	-1	0	0	1	0	a ₀	0.61
v ₋₁₋₁		1	-1	1	-1	1	1	a ₁	0.72
v ₀₋₁		1	0	-1	0	0	1	a ₂	0.68
v ₀₋₀		1	0	0	0	0	0	a ₃	0.67
v ₀₋₁		1	0	1	0	0	1	a ₄	0.76
v ₁₋₁		1	1	-1	-1	1	1	a ₅	0.74
v ₁₋₀		1	1	0	0	1	0		0.61
v ₁₋₁		1	1	1	1	1	1		0.73
									0.69

Solution – improved position of the best fit

	local system	search area	standard deviations [pel]
r _{max} [pel]	0.05	30.05	$\sigma_r = 0.01$
c _{max} [pel]	0.22	32.22	$\sigma_c = 0.01$

Fig. B1 Calculation of an improved position of the best fit of a réseau cross from Fig. 1.10.

The values of standard deviation are probably too optimistic due to a good agreement between derived polynomial and calculated values of correlation coefficient. With increasing

number of observations (i.e. a polynomial is calculated over a larger area), a fit between a polynomial and an original discrete correlation function decreases and standard deviations of a determined position get worse as Tab. B.1 shows. In the presented example all the achieved values of standard deviations are relatively small. However the tendency is visible. For the area $7 \times 7 \text{ pel}^2$ and larger, the standard deviation exceeds the value of the shift in the row direction which signalizes low reliability of a found position.

Number of observations	Subpixel position of the best fit within a search area		Standard deviation	
	row	col	σ_r [pel]	σ_c [pel]
3x3	30.05	32.22	0.01	0.01
5x5	30.03	32.21	0.02	0.02
7x7	30.03	32.16	0.05	0.05
9x9	30.02	32.13	0.09	0.08

Tab. B.1 Improved position of the best fit of the cross from Fig. 1.10 calculated with different a number of observations in the adjustment of polynomial.

Choosing an optimal number of observations is a task depending individually on a shape of a specific discrete correlation function. In the given example the correlation function is relatively smooth and contains only one peak in a small area close to the maximum of correlation coefficient. Therefore a window of $3 \times 3 \text{ pel}^2$ or $5 \times 5 \text{ pel}^2$ seems to be optimal for fitting a 2nd order polynomial. The redundancy number of k , ($k=m-6$, $m=9$, 15) gives a possibility of excluding possible outliers in least squares adjustment by means of data snooping or robust adjustment (see chapter 1.3).

By combining MATLAB functions *correl_coef.m* and *c_subpixel.m*, an improved position of the best fit together with standard deviations of subpixel shifts can be obtained for an input template and search area. Detection of outliers is not included. The number of observations in a subpixel calculation is not fixed to e.g. $3 \times 3 \text{ pel}^2$ but depends on a size of ‘correlation coefficient matrix’ (see upper left part of Fig. B1) that is an input of a function *c_subpixel.m*. Both functions can be found on the attached CD.

B.2 Least squares matching

In least squares matching (LSM), geometric and radiometric transformation parameters between two image patches are to be found by means of minimising differences between their grey values. Knowing the geometric transformation parameters, a position of a centre of one patch can be found within the second one with subpixel accuracy. In the following text, general formulas of LSM are described in details for image patches \mathbf{g}_1 and \mathbf{g}_2 . An explanation of differences in resampling procedures when a design matrix of least squares adjustment is derived according to a search or a template window (formulas 1.7 and 1.8, chapter 1.2.1.2) is also included.

An observation equation of LSM between two image patches \mathbf{g}_1 and \mathbf{g}_2 is given by formula B.7.

$$g_1(r, c) + v(r, c) = g_2(f_r(p_1, \dots, p_n, r, c), f_c(p_1, \dots, p_n, r, c)) r_s + r_t = \bar{g}_2(r, c) \quad (B.7)$$

$f_r, f_c \dots$ functions representing geometric transformation between image patches

$p_i \dots \dots \dots$ geometric parameters

$r_t, r_s \dots \dots \dots$ radiometric parameters (shift and scale)

$\bar{g}_2 \dots \dots \dots$ adjusted grey values in the image patch 2

$\mathbf{v} \dots \dots \dots$ vector of residuals in grey values

In order to calculate both geometric and radiometric parameters in least squares adjustment, the right side of the equation B.7 has to be linearized with respect to the unknowns p_1, \dots, p_n, r_s , and r_t as formula B.8 shows. The linear radiometric parameters are approximated as $r_s = r_s^0 + dr_s = 1 + dr_s$ and $r_t = r_t^0 + dr_t = 0 + dr_t$.

$$g_1(r, c) + v(r, c) = \bar{g}_2^0(r, c) + \frac{\partial \bar{g}_2(r, c)}{\partial g_2} \left(\frac{\partial g_2(r, c)}{\partial f_r} df_r + \frac{\partial g_2(r, c)}{\partial f_c} df_c \right) + \frac{\partial \bar{g}_2(r, c)}{\partial r_s} dr_s + \frac{\partial \bar{g}_2(r, c)}{\partial r_t} dr_t$$

$$g_T(r, c) + v(r, c) = \bar{g}_2^0(r, c) + (g_{2R}(r, c) df_r + g_{2C}(r, c) df_c) r_s^0 + \bar{g}_2^0(r, c) dr_s + dr_t$$

for $r_s^0 = 1$

$$g_T(r, c) + v(r, c) = \bar{g}_2^0(r, c) + g_{2R}(r, c) df_r + g_{2C}(r, c) df_c + \bar{g}_2^0(r, c) dr_s + dr_t \quad (B.8)$$

$$\bar{g}_2^0(r, c) = g_2(f_r^0(r, c), f_c^0(r, c)) r_s^0 + r_t^0$$

$$df_r(r,c) = \frac{\partial f_r(r,c)}{\partial p_1} dp_1 + \dots + \frac{\partial f_r(r,c)}{\partial p_n} dp_n$$

$$df_c(r,c) = \frac{\partial f_c(r,c)}{\partial p_1} dp_1 + \dots + \frac{\partial f_c(r,c)}{\partial p_n} dp_n$$

$g_{2R}(r,c), g_{2C}(r,c) \dots$ gradients in grey values in row and column directions

$\bar{g}_2^0(r,c) \dots$ grey values of patch 2 after applying approximations of both geometric and radiometric parameters

Gradients in grey values or slopes in densities are calculated in their simplest form as $g_{2R}(r,c) = (g_2(r+1,c) - g_2(r-1,c)) / (2pel)$, $g_{2C}(r,c) = (g_2(r,c+1) - g_2(r,c-1)) / (2pel)$ where $r=1, \dots, R$, $c=1, \dots, C$ and pel stands for the pixel size ($pel=1$ when working in the pixel co-ordinate system). At the edge of the window the gradients are calculated from the grey values of an evaluated pixel and its nearest neighbour in a row or column direction, i.e. $g_{2R}(1,c) = (g_2(2,c) - g_2(1,c)) / pel$. Gradient operators like Sobbel or Prewitt operator can also be applied. In contrary to the presented calculation of gradients they average values from six surrounding pixels and therefore are more resistant to noise (Mikhail et al., 2001).

An overview of geometrical models used in least squares matching (LSM) is given in Tab. 1.3, chapter 1.2.1.2. Depending on the geometrical model, the differential relations $df_r(r,c)$ and $df_c(r,c)$ can be expressed in forms shown in Tab. B2.

Transformation	Transformation equation	Differentiation of transformation equation
Conform – 2 parameters t_r, t_c	$f_r(r,c) = r + t_r$ $f_c(r,c) = c + t_c$	$df_r(r,c) = dt_r$ $df_c(r,c) = dt_c$
Conform – 3 parameters t_r, t_c, k	$f_r(r,c) = kr + t_r$ $f_c(r,c) = kc + t_c$	$df_r(r,c) = rdk + dt_r$ $df_c(r,c) = cdk + dt_c$
Conform – 2 parameters t_r, t_c, k, α *)	$f_r(r,c) = ar + bc + t_r$ $f_c(r,c) = -br + ac + t_c$	$df_r(r,c) = rda + cdb + dt_r$ $df_c(r,c) = -rdb + cda + dt_c$
Affine - 6 parameters $t_r, t_c, k_r, k_c, \alpha, \kappa$ **)	$f_r(r,c) = a_1 r + a_2 c + t_r$ $f_c(r,c) = b_1 r + b_2 c + t_c$	$df_r(r,c) = rda_1 + cda_2 + dt_r$ $df_c(r,c) = rdb_1 + cdb_2 + dt_c$

*) $a = k \cos \alpha, b = k \sin \alpha$

**) $a_1 = k_r \cos \alpha, a_2 = k_r \sin \alpha, b_1 = k_c \sin(\kappa - \alpha), b_2 = k_c \cos(\kappa - \alpha)$

Tab. B.2 Differentiation of geometric transformation equations used in least squares matching. t_r and t_c are translation parameters, k respectively k_r and k_c stand for scale factors, α for rotation and κ for orthogonality.

In a matrix notation linearized observation equations B.8 can be written as

$$\mathbf{v} = \mathbf{A} \mathbf{dx} - \mathbf{l} \tag{B.9}$$

\mathbf{dx} vector of unknown geometric and radiometric parameters

\mathbf{A} design matrix containing derivatives of the function g_2 with respect to all unknown parameters

\mathbf{l} vector of differences between grey values in the image patches with RC elements
 $\mathbf{l} = \{g_1(r,c) - g_2^0(r,c)\}, r=1, \dots, R, c=1, \dots, C$

\mathbf{v} vector of residuals

In case of affine transformation and two radiometric parameters, $\mathbf{dx} = [da_1 \ da_2 \ dt_r \ db_1 \ db_2 \ dt_c \ dr_r \ dr_c]^T$ and \mathbf{A} is a $RC \times 8$ matrix with elements $\mathbf{A} = \{rg_{2R}(r,c) \ cg_{2R}(r,c) \ g_{2R}(r,c) \ rg_{2C}(r,c) \ cg_{2C}(r,c) \ g_{2C}(r,c) \ g_2 \ 1\}, r=1, \dots, R, c=1, \dots, C$. If a less-parameters geometric model is applied or a radiometric adjustment is done prior to LSM, the number unknowns and the number of columns in the \mathbf{A} matrix decrease accordingly.

The condition of least squares leads to a solution that was already described in previous section by equation B.3. A weight matrix \mathbf{P} is an identity matrix at the beginning of calculation. It can be changed in order to weight down observations with high residuals, e.g. pixels which grey values do not fit to an applied geometric and radiometric model.

As mentioned in the chapter 1.2.1.2, after each iteration the search area has to be resampled. A new search window of the same size as a template window is created. The grey values for each pixel must be interpolated from an original image patch. Depending on whether the transformation parameters were derived for a template or search window (formula 1.8 or 1.7 chapter 1.2), direct or indirect transformation is used as Fig. B2 shows.

The results of the least squares matching are transformed pixel co-ordinates of the centre of the template in the search area (formula B.10).

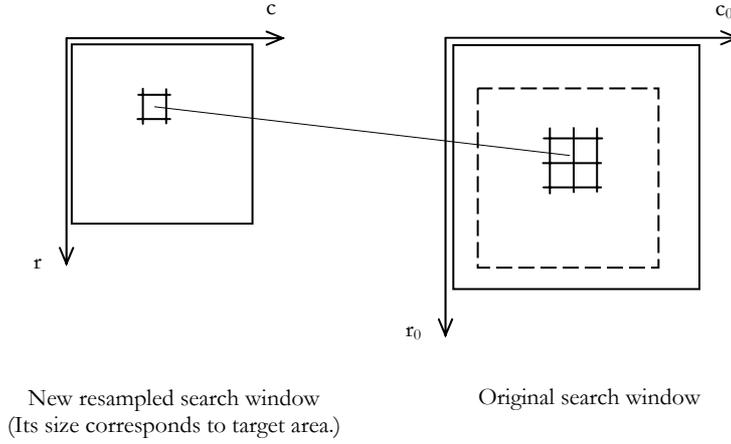
$$\begin{aligned} \bar{r} &= f_r(\bar{p}_1, \dots, \bar{p}_n, r_C, c_C) \\ \bar{c} &= f_c(\bar{p}_1, \dots, \bar{p}_n, r_C, c_C) \end{aligned} \tag{B.10}$$

\bar{r}, \bar{c} position of the best fit of the centre of the image patch 1 in the image patch 2

$\bar{p}_1, \dots, \bar{p}_n$... adjusted geometric transformation parameters

r_C, c_C co-ordinates of the centre of the image patch number 1

Co-ordinates r_C, c_C are equal to a half of the image patch size rounded to the nearest integer value towards infinity (in case that the image patch has odd number of rows and columns).



a) Observation equation:

$$g_T(r_0, c_0) = g_S(a_1 r_0 + a_2 c_0 + t_c, b_1 r_0 + b_2 c_0 + t_c) r_s + r_t$$

The grey values in a new search window are interpolated at the position

$$\begin{pmatrix} r_0 \\ c_0 \end{pmatrix} = \begin{pmatrix} a_1 & a_2 \\ b_1 & b_2 \end{pmatrix} \begin{pmatrix} r \\ c \end{pmatrix} + \begin{pmatrix} t_r \\ t_c \end{pmatrix}$$

Radiometric transformation $g_S, r_s + r_t$

b) Observation equation:

$$g_S(r, c) = g_T(a_1 r + a_2 c + t_r, b_1 r + b_2 c + t_c) r_s + r_t$$

The grey values in a new search window are interpolated at the position

$$\begin{pmatrix} r_0 \\ c_0 \end{pmatrix} = \frac{1}{a_1 b_2 - a_2 b_1} \begin{pmatrix} b_2 & -a_2 \\ -b_1 & a_1 \end{pmatrix} \begin{pmatrix} r - t_r \\ c - t_c \end{pmatrix}$$

Radiometric transformation $(g_S - r_t) / r_s$

Fig. B2 Differences in resampling procedures of a search image patch if the transformation parameters are calculated with respect to the search window (a) and the template window (b). The advantage of case b) is that the design matrix does not change during the whole iterative process.

Based on the law of error propagation, the variances in transformed co-ordinates of the template centre into the search area are calculated by formula B.11:

$$\begin{bmatrix} \sigma_r^2 & \sigma_c^2 \end{bmatrix} = \text{diag}(\Sigma_{RC}) \quad (B.11)$$

$$\Sigma_{RC} = \mathbf{B} \Sigma_x \mathbf{B}^T$$

Σ_{RC} ... covariance matrix

Σ_x ... covariance matrix of adjusted geometric parameters (see formula B.5)

$$\mathbf{B} = \begin{bmatrix} \frac{\partial \bar{r}}{\partial \bar{p}_1} & \cdots & \frac{\partial \bar{r}}{\partial \bar{p}_n} \\ \frac{\partial \bar{c}}{\partial \bar{p}_1} & \cdots & \frac{\partial \bar{c}}{\partial \bar{p}_n} \end{bmatrix}$$

In (Atkinson, 1996), it is suggested to evaluate a quality of the match only by means of standard deviations of the shift parameters σ_{tr} and σ_{tc} given by formula B.12:

$$\begin{aligned}\sigma_{tr} &= \sigma_0 \sqrt{Q_{ii}} \\ \sigma_{tc} &= \sigma_0 \sqrt{Q_{jj}}\end{aligned}\tag{B.12}$$

σ_0 standard deviation of a unit weight

Q_{ii}, Q_{jj} diagonal elements of a cofactor matrix $\mathbf{Q} = \mathbf{N}^{-1}$ (see formula B.5) corresponding to the shift parameters

Whether two unknown geometric and radiometric parameters are correlated can be seen from the correlation matrix ρ . Correlation coefficients r_{ij} are calculated from the elements of a covariance matrix Σ_x (see equation B.5) by means of formula (B.13). Both matrices Σ_x and ρ are symmetric. If a correlation coefficient between two unknowns is in its absolute value close to 1, the used geometric and radiometric model should be re-evaluated and probably simplified and a new calculation carried out. Similarly, if a standard deviation of an unknown exceeds a value of a calculated parameter, the LSM should be recalculated with a new transformation model that does not contain that unreliable parameter.

$$\Sigma_x = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \dots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \dots & \sigma_n^2 \end{pmatrix} \quad \rho = \begin{pmatrix} 1 & r_{12} & \dots & r_{1n} \\ r_{21} & 1 & \dots & r_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{n1} & r_{n2} & \dots & 1 \end{pmatrix} \quad r_{ij} = \frac{\sigma_{ij}}{\sigma_i \sigma_j}\tag{B.13}$$

Σ_x ... covariance matrix

ρ ... correlation matrix

r_{ij} ... correlation coefficient

The program solution of least squares matching used in this project is divided into three MATLAB functions that can be found in Appendix C:

1. The main function **least_sq.m** calculates improvements of the position of the best fit of the centre of the template within a search area and its standard deviation in an iterative process. The input parameters are a template matrix, a search area matrix and the approximate position of the best fit obtained by means of correlation coefficient. In each iteration functions **LSM.m** and **resampling.m** are called.

2. Function ***LSM.m*** includes least squares adjustment. The input parameters are a template matrix and a resampled section of a search matrix. It returns corrections of transformation parameters, a standard deviation of a unit weight and a cofactor matrix.
3. Function ***resampling.m*** returns a resampled section of a search matrix. The input parameters are a search matrix, transformation parameters, the size of the template and its position within a search area.

In order to carry out all the tests described in chapters 1.2.1.5, 2.4.2, and 3.3.2 all three functions were designed for three different geometric models (translation, conform, and affine) both with and without radiometric parameters and both for a stable and recomputed design matrix A . In case of not treating radiometric parameters as unknowns, the image patch used for derivation a design matrix is radiometrically corrected. The function ***lin_rad_adj.m*** calculates radiometric shift r_t and scale r_s between the image patches by means of least squares adjustment.