VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ
BRNO UNIVERSITY OF TECHNOLOGY

FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ
ÚSTAV RADIOELEKTRONIKY

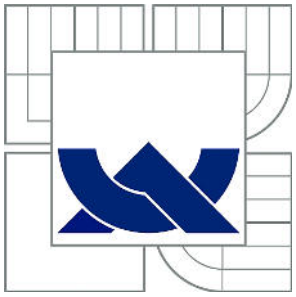
FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION
DEPARTMENT OF RADIO ELECTRONICS

SELECTED PROBLEMS IN PHOTOGRAMMETRIC SYSTEMS ANALYSIS

DIZERTAČNÍ PRÁCE
DOCTORAL THESIS

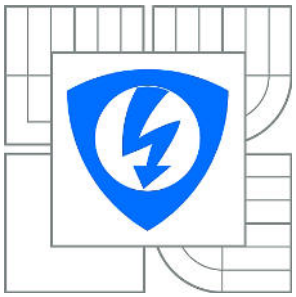
AUTOR PRÁCE Ing. LIBOR BOLEČEK
AUTHOR

BRNO 2014



VYSOKÉ UČENÍ TECHNICKÉ V BRNĚ

BRNO UNIVERSITY OF TECHNOLOGY



FAKULTA ELEKTROTECHNIKY A KOMUNIKAČNÍCH TECHNOLOGIÍ

ÚSTAV RADIOELEKTRONIKY

FACULTY OF ELECTRICAL ENGINEERING AND COMMUNICATION

DEPARTMENT OF RADIO ELECTRONICS

SELECTED PROBLEMS IN PHOTOGRAMMETRIC SYSTEMS ANALYSIS

VYBRANÉ PROBLÉMY ANALÝZY FOTOGRAMMETRICKÝCH SYSTÉMŮ

DIZERTAČNÍ PRÁCE

DOCTORAL THESIS

AUTOR PRÁCE — Ing. LIBOR BOLEČEK

AUTHOR

VEDOUCÍ PRÁCE — prof. Ing. VÁCLAV ŘÍČNÝ, CSc.

SUPERVISOR

BRNO 2014

ABSTRACT

This dissertation deals with selected topics of digital photogrammetry. The problem is defined and the state of the art is described in the first part of the dissertation. Four specified aims are solved. The proposal of the method for finding corresponding points is the first topic. Two new methods were proposed. The first method uses conversion of an image to pseudo- colors. The second method used a probabilistic model obtained from the known pairs of the corresponding points. The analysis of the accuracy of the reconstruction is the second solved topic. The influence of the various aspects to the accuracy of the reconstruction is analyzed. The most attention is paid to incorrect camera alignment and errors in finding corresponding points. The third topic is estimation of the depth maps. The two method were proposed. The first method is based on the combination of the passive and active method. The second wholly passive approach uses continuity of the depth map. The last investigative topic is quality of experience of the 3D videos. The subjective tests of the perception of 3D content for the various 3D displaying systems were performed. The dependency of the perception on the viewing angle was also investigated.

KEYWORDS

digital photogrammetry, corresponding points, depth map, QoE subjective test,

ABSTRAKT

Disertační práce se zabývá vybranými partiemi digitální fotogrammetrie. V první části práce je definované téma a popsán současný stav poznání. V následujících kapitolách jsou postupně řešeny čtyři dílčí navzájem navazující cíle. První oblastí je návrh metody pro hledání souhlasných bodů v obraze. Byly navrženy dvě nové metody. První z nich používá konverzi snímků do nepravých barev a druhá využívá pravděpodobností model získaný ze známých párů souhlasných bodů. Druhým tématem je analýza přesnosti výsledné rekonstrukce prostorových bodů. Postupně je analyzován vliv různých faktorů na přesnost rekonstrukce. Stěžejní oblastí je zkoumání vlivu chybného zarovnání kamer a chyby v určení souhlasných bodů. Třetím tématem je tvorba hloubkových map. Byly navrženy dva postupy. První přístup spočívá v kombinaci pasivní a aktivní metody druhý přístup vychází z pasivní metody a využívá spojitosti hloubkové mapy. Poslední zvolenou oblastí zájmu je hodnocení kvality 3D videa. Byly provedeny a statisticky vyhodnoceny subjektvní testy 3D vjemu pro různé zobrazovací systémy v závislosti na úhlu pozorování.

KLÍČOVÁ SLOVA

digitální fotogrammetrie, souhlasné body, hloubková mapa, subjektivní testy QoE ,

BOLEČEK, Libor *Vybrané problémy analýzy digitálních fotometrických systémů*: doctoral thesis. Brno: Brno University of Technology, Faculty of Electrical Engineering and Communication, Ústav radioelektroniky , 2014. 136 p. Supervised by prof. Ing. Václav Říčný, CSc.

DECLARATION

I declare that I have written my doctoral thesis on the theme of "Vybrané problémy analýzy digitálních fotometrických systémů" independently, under the guidance of the doctoral thesis supervisor and using the technical literature and other sources of information which are all quoted in the thesis and detailed in the list of literature at the end of the thesis.

As the author of the doctoral thesis I furthermore declare that, as regards the creation of this doctoral thesis, I have not infringed any copyright. In particular, I have not unlawfully encroached on anyone's personal and/or ownership rights and I am fully aware of the consequences in the case of breaking Regulation § 11 and the following of the Copyright Act No 121/2000 Sb., and of the rights related to intellectual property right and changes in some Acts (Intellectual Property Act) and formulated in later regulations, inclusive of the possible consequences resulting from the provisions of Criminal Act No 40/2009 Sb., Section 2, Head VI, Part 4.

Brno ....

(author's signature)

ACKNOWLEDGEMENT

I would like to thank Prof. Ing. Vaclav Ricny , CSc. for mentoring, consultation, patience and inspiring suggestions to work.

Brno . . . . . .

. . . . . .

(author's signature)

CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 INTRODUCTION

1.1 Problem formulation

Photogrammetry is a scientific discipline which deals with reconstructing three-dimensional objects from two-dimensional photographs. Photogrammetry allows the reconstruction of an object and analysis of its characteristics without physical contact with them [81]. This work deals with multiple frame photogrammetry allowing to determine three spatial coordinates. Designation stereogrammetry is used in the situation when two cameras with parallel optical axes are used and their spatial positions differ only in horizontal direction. The first use of analog photogrammetry is dated approximately to the second half of the 19th century. At this time, a mathematical element was defined (see section 2.1). The input to photogrammetry are two or more photo images acquired by a camera system. All principles discovered and described during the origin of this method are still in force. The new age of photogrammetry began after the coming of digital photography. Digital photogrammetry uses a digital camera to obtain an image of the scene. Subsequently, a personal PC is used for data processing. In recent years, digital photogrammetry and stereophotogrammetry have become a dynamically developing scientific area. This fact relates with rapid expansion of 3D technology. We can observe the development of new sensing and displaying systems. This development was influenced by increasing performance of computers. The performance of current PCs allows to execute computationally difficult operations. Due to this fact, many operations can be executed even in real time.

The use of computers, special programs and modern technical devices allow the implementation of basic algorithms needed for solving classic stereophotogrammetry as image orientation and triangulation. Moreover, new approaches and methods can be used. Current digital photogrammetry includes image processing. Photogrammetry can be divided into three categories: analogue photogrammetry, analytic photogrammetry using PCs and digital photogrammetry which includes machine vision, computer vision and pattern recognition. Nowadays, basic algorithms for analyzing and solving fundamental tasks are known and many are published in expert literature, nevertheless, areas and topics, in which finding new methods are required, still occur.

Image processing is used in every step of obtaining information about spatial coordinates. In preprocessing, the following methods are applied: filtration, change of contrast, sharpening and some others. Appropriate preprocessing of the input images brings better and reliable results. Finding corresponding points is a crucial method used during proper reconstruction. Methods for image segmentation are

used in the course of estimating depth maps. Artificial intelligence and optimization methods are frequently used in model reconstruction.

Reconstructing spatial coordinates and obtaining depth maps is a problem, which have wide applications in many areas. The depth of a point can be represented in two various ways. The first way is creating a model with discrete points $P(X, Y, Z)$. These points are described by three spatial coordinates (X,Y,Z). The coordinates are in a certain coordinate system with determined origin. The second way is an expression using a depth map. The depth map is an image with the same size as the input image. The value of the individual pixel of the depth map is given by the relative depth of the scene, where depth means the distance of the point from the camera. This value is equal to the spatial coordinate Z.

Civil and mechanical engineering industries are typical representatives of the fields which use a 3D model of the scene. Other disciplines using spatial reconstruction are, for example, medicine, robotics, reconstruction of traffic accidents and the entertainment industry. This reconstruction can be utilized for modeling buildings, rooms or various objects. Subsequently, reconstructed models can be used during the proposed building process, reconstruction of the building or during the creation and testing of some instruments. Nowadays, 3D TV is becoming more popular and more used. The first standards about 3D TV have already been formed.

The basis of stereoscopic displays is using two images of the same scene. These images are called stereo images (left and right). Each eye sees a slightly different image of the same scene. The images are shifted in a horizontal direction. The resulting spatial image is formed in the brain. The depth map plays an important role during the creation of stereo images and transmission of the data for 3D imaging. The quality of the depth map is of fundamental importance for spatial perception.

The main topic and aim of my dissertation is obtaining information of spatial coordinates by using two digital images. The dissertation deals with both representation of spatial information (depth map and spatial coordinates). The first section deals with metric reconstruction of the spatial model and it contains a proposal of the new approach for finding corresponding points in the images. Another part of the dissertation also analyses the impact of various aspects on the accuracy of the reconstruction. The dissertation also deals with the estimation of depth maps. The last part of the dissertation deals with the quality of experience in 3D video.

1.2 State of the art

1.2.1 Photogrammetry

The basic conception of photogrammetry has been known for a relatively long time. Aimé Laussedat with Albrecht Meydenbauer are regarded as founders of photogrammetry. Meydenbauer firstly used the word photogrammetry in his paper in 1893. Current research uses the basic conception and has expanded this foundation. The principal trend of the current progression is to obtain a more accurate model and faster execution of each step of the 3D model reconstruction. The process of the reconstruction can be divided into three essential steps. Finding the corresponding points is the first of them. The corresponding points are image points which represent the same spatial point in the scene. The second task is calibration of the used cameras. Camera calibration follows, after finding the corresponding points. Camera calibration builds on the information about the position of the corresponding points. An effort is made so that these operations are done automatically and effectively. A more detailed description of these steps follows in the next statements.

Camera calibration can be divided into interior calibration and exterior calibration. The interior calibration is executed for the purpose of finding camera parameters. The interior parameters are focus distance f, position of the principal point (x_{0,y_0}) and distortion d_s. The output of the exterior calibration is information about the mutual position of the two cameras. The information about their positions is represented by rotation matrix $\mathbf{R}$ and translation vector $\mathbf{T}$. These parameters can be extracted from the essential matrix. The essential matrix can be calculated based on knowing the minimum of seven pairs of corresponding points. The essential matrix is a special form of the fundamental matrix.

The methods for calculating interior parameters of the camera can be divided into two groups:

- off-line calibration,
- self-calibration.

The methods from the first group are executed using certain calibration patterns of the regular shape (frequently chessboard) with known properties. This calibration pattern is captured from various positions:

- various viewing (captured) angles ,
- various positions of the cameras.

Subsequently, calibration is performed by finding significant points in the scene and evaluating the change of their positions. Off-line calibration is executed before capturing the reconstructed scene. The calibration matrix $\mathbf{K}$ is calculated and used in the subsequent reconstruction of the scene. The representative of this group can

be the method published by Z. Zhang [1]. The output of this method is very reliable. However, these methods have some disadvantages. The first disadvantage is the requirement of the calibration pattern. The second disadvantage is the impossibility to react to the camera focusing on an object which changes camera calibration.

The self-calibration methods do not require a calibration pattern. These methods are executed directly using images of the reconstructed scene. The first mention of this approach can be found in [2]. The authors of this method called it automatic calibration. Fauregas [3] and Hartley [4] proved that we can obtain projective reconstruction from two uncalibrated images, even without knowing the camera calibration. Until now, many algorithms have been proposed. The research from Kruppa was basic for creating a whole group of algorithms for camera calibration [3],[5],[6],[7] and also for algorithms for robust and accurate estimation of the fundamental matrix **F** [8]-[11]. In some methods, calibration is executed in one step. However, other methods use the stratified approach. Stratification to three phases (projective, affine, euclidean) was first published in [12]. Subsequently, this approach was used for formulating a mathematical system for spatial reconstruction [13]. However, Euclid reconstruction is not always feasible using the stratified approach. The interior ambiguity in reconstruction were studied in [14],[15],[16]. Calibration can also be executed using vanishing points. The methods for detecting vanishing points and calculation of vanishing lines can be found for example in [17] , [18].

Finding image correspondence is a very important process for reconstructing spatial coordinates. Image correspondence is determined using the determination of the corresponding points in both images. Finding corresponding points is a frequently examined topic. The process of finding corresponding points has two steps: finding significant points in both images and subsequently creating pairs from the same image points (matching). The correspondence issue was first solved in the 70's in the last century. Marr a Poggio introduced a detailed theory of stereo vision [19]. In this work, the ambiguity of the correspondence problem was defined. The authors introduced the concept of continuity. Moreover, two basic constraints for the definition of matching points are described. The same authors proposed a method based on matching salient features. In this moment, the concept of feature points arose. The edges or the corners can be salient features. The principle was subsequently improved by Pollard, Mayhew and Frisb [20]. In subsequent years, researchers dealt with modifying the basic theory. The nascent methods were divided into two basic groups. The global methods are the first group. More about global methods can be found in [21],[22], [23]. Dynamic programming is closely related with global methods. The local methods are the next possible approach. Local methods compare individual points for matching. The essence of local methods consists in determining the degree of similarity. However, we compare a suitable neighborhood (window) of

the examined point not only itself point. The following parameters can be used for comparing SSD (Sum of Squared Differences), SAD (Sum of Absolute Differences), NCC (Normalized cross-correlation) [24]. etc. One of the methods based on SAD is proposed by Hamzah [25].

Many algorithms were proposed for solving this issue. A survey of the proposed methods can be found in articles [26], [27], [28]. The authors generally divide the methods into a few groups: methods based on contours, methods based on the intensity and methods based on the parametric model. The Harris detector [31] is a frequently used algorithm. Currently, the most used methods are so called descriptors. The descriptors provide some description of the found significant points. The description is subsequently used for matching. The most frequently used descriptors are Scale Invariant Feature Transform (SIFT) [30] and its improvement Speeded-Up Robust Feature (SURF) [29]. A high quality detector should be invariant against various image adjustments, for example: rotation, translation, scaling, adding noise, and changing illumination. The detectors, even descriptors, work with an image in greyscale in most cases.

1.2.2 Generation of the depth map

Methods for estimating depth maps can be divided into two groups: active methods and passive methods. The active methods can be denoted as active scanning. The feature of these methods is using extra information. The extra information is often added using a projection of the same light [32],[33]. The active method using projection can be further divided into coherent or incoherent dependency of the type of light used. The method published in [34] can be representative of coherent methods. The incoherent method can use a projection of the fringe pattern. The method using a projection of the fringe pattern was first described and used in [35] , [36], [37]. This approach of estimating the depth map can be described as a conversion change of phase to depth. Some various variants of this method exist. The first one is based on filtering the frequency domain [38], called Fourier Transform Profilometry FTP. Filtration in the time domain can also be used [39]. These two variants utilize only one projection. On the contrary, the method called Phase-Shifting [40] is based on the repeated projection of the pattern with various initial phases. The summary of this method can be found in [112].

This dissertation is mostly devoted to the passive method. The passive methods utilize the procedure dense matching. Dense matching extends correspondences of individual points to the correspondences of segments. The segments are obtained using a segmentation technique. These methods frequently use tools of artificial intelligence, optimization or dynamic programming. Therefore, the important step

of estimating the depth map by this approach is finding the corresponding points. If two images in normal form are available, then corresponding points lie in the same row. In other words, the positions of corresponding points differ only in horizontal coordinates. This horizontal difference is called horizontal parallax. Subsequently, information about depth is given by horizontal parallax. The depth increases with decreasing horizontal parallax. Therefore, there is an indirect proportion between depth and horizontal parallax. The input images can have a general relationship (corresponding points are not in the same row). Then, we have two possibilities how to obtain the depth map. Converting images to the normal case is one of the possible approaches. This operation is called image rectification [42], [43]. Rectification is executed by finding the corresponding points. Subsequently, transforming the images to common space is executed. The second way is estimating the depth map based on the depth of the reconstructed point obtained using stereophotogrammetry (see chapter 2.1). An extensive survey of the passive method for estimating the depth map can be found in [44], [45].

1.2.3 Quality evaluation and accuracy of the reconstruction

The influence of various factors on the accuracy of the reconstruction was the subject of many studies of various authors. Kytö, Nuutinen and Oittinen in [46] examined the influence of the change in the stereo base and focal distance on the depth resolution. The study assumed accurate determination of these parameters. The authors compared the theoretical resolution ability of a man with achievable resolution for capturing an image with a given stereo base and focal distance. The achievable results show that a change of the stereo base has a greater influence than a change in the focal distance. Accuracy increases with increasing stereo base and focal distance. Therefore the optimal stereo base is infinitely large, however, this idea has many advantages. This assumption refutes Zhang in article [48]. The article deals with analyzing error of the reconstruction in dependency on the stereo base and on mutual camera rotation. The author designed an error model for one, even two, cameras. The author also mentioned that standardized methods for error evaluation did not exist in their days. Zhao and Nandhakumar [49] deal with the influence of inaccurate determination of parameters of exterior calibration. This inaccuracy can be represented in two ways: error in the rotation matrix and translation vector or by error angles or translation between cameras. The article contains the analysis of the influence of each parameter separately. Some authors deal with the influence of image discretization, hence the consequence of the finite size of the pixel [50]. Using the error due to discretization, it is possible to determine the tolerance error in the calibration parameters [51]. Accurate determination of the corresponding points

has a fundamental importance in the overall accuracy. The accuracy of the spatial coordinates generally decreases quadratically with increasing depth. The authors of the publication [52] tried to solve this problem. They proposed a system with a variable stereo base. The proposed system has a constant error for variable depth. Article [53] deals with errors related to finding corresponding points. The authors found errors arising in edge detection during the process of finding significant points. Subsequently, they examined the spreading of this error during the whole process of reconstruction. The authors considered three sources of errors:

- inaccuracy in the camera model,
- inaccuracy in exterior calibration,
- inaccuracy arising during image processing.

The definition of these three sources of error can be found in their papers [54] and [55]. In these articles, the authors considered the importance of the accuracy of finding corresponding points, however, they not quantitative.

3D television systems have become popular and different systems for 3D imaging are used today. In consequence, a lot of research is devoted to this topic. The quality of the experience (QoE) in 3D video is an actual topic. The evaluation of the quality and spatial effect of the 3D videosequences is a very complicated problem. The evaluation can be subjective or objective. In the last few years, intensive research has been executed in this area. This issue has a lot in common with evaluating virtual reality [56]. The proper definition of QoE appeared in Qualinet White Paper [57] on Definitions of Quality of Experience created by the European Network on quality of experience in multimedia systems and services. The research on QoE in 3D has a few aims. In the beginning, researchers wanted to discover new requirements in the area of QoE [58], [59]. A subsequent aim was to examine the possibility of using the current method used originally for 2D video [60], [61], [62], [63]. However, the requirement on new objective metrics is obvious. A tool for its proposal is by executing subjective tests. ITU recommendations [64], [65], [66] specified conditions for subjective tests. Therefore many teams are carrying out subjective tests and examining dependency of QoE on various aspects. A number of objective metrics have been proposed in recent years. One group used a depth map for estimating the quality [67],[68]. Authors lead by X. Liyuan deal with the impact of cross-talks [69], [70]. The same work group considered in [71] the influence of scene content, camera baseline and screen size. Another important topic in QoE in 3D is examining the influence of coding and quality for a synthetic image [72], [73], [74], [75]. In this area, authors investigate the dependency of the result of the subjective test on the used method for coding and depth map rendering. Concurrently, they examine the correlation of the results with the objective metrics. The formation of new artifacts during depth map rendering is the fundamental reason for this research. Some of

the authors deal with the impact of the parameters of the camera and capturing conditions on the stereo effect [76], [77], [78]. Yamanoue deals with the influence of TV display parameters [79].

1.3 Aim of the work

The aims of my dissertation were established after studying scientific literature and analyses of the state of the art. Notwithstanding many detectors of the significant points and algorithms for finding corresponding points, there are issues which need to be solved. One problem is finding corresponding points for a specific image point belonging to an area without contrast or with regular texture. The analyzing of the accuracy of the reconstruction logical follows on the proposal the method for finding corresponding points affecting accuracy. Estimating the depth map and evaluating the spatial effect of the watcher is a very perspective area at this time of dynamic development of commercial 3D imaging. The solution of described issues can be summarized to the following aims.

1. Proposal of novel fast methods for matching points in images. The proposal will be supported by an analysis of the currently used methods. Software implementation of the proposed methods to the system for determining spatial coordinates of the points in the scene.
2. Analysis of the achievable accuracy of determining spatial coordinates in the 3D model of the scene. The quantification of the aspects affecting accuracy (especially the aspects related with parameters of the sensing system and its calibration).
3. Proposal of the system (algorithm)for estimating the depth map from the two images of the scene.
4. The realization and evaluation of the subjective tests of the spatial effect and quality in a 3D TV. The examination of the dependency of spatial perception on the various parameters: sensing parameters, content of the sequence, viewing conditions during reproduction on TV displays using various 3D systems.

2 3D METRIC RECONSTRUCTION

The spatial metric reconstruction of the spatial model of an analyzed scene using photogrammetry (location of the spatial coordinates) is the central topic of my dissertation. The elementary requirement for using photogrammetry is the acquirement of a minimum of two photos of an analyzed scene. The fundamental ideas of photogrammetry are described for example in books [80], [81] and [82]. Mathematical apparatuses for this reconstruction is described in chapter 2.1. The fundamental task in this reconstruction is finding corresponding points in both input images of the analyzed scene. Chapter 2.2 contains a comparison of some methods frequently used for finding corresponding points. Finding corresponding points is a great problem especially in automatic and semi-automatic systems. A high quality solution of this problem is very difficult especially in an image area with the following properties:

- regular texture,
- small brightness variation.

Two novel approaches for finding corresponding points are proposed in this chapter. The first proposed method is based on the presumption that if the depth of a few points in the neighborhood of the selected point is known, then the depth of the selected point can be determined without directly finding its corresponding point. This method is primary designed to reconstruct individual points selected by the user. Converting gray scale images to pseudo-colors is another new proposed approach to the problem of finding corresponding points. The proposed methods are described in sections 2.3 and 2.4.

2.1 Reconstruction of the spatial model from two uncalibrated images

The mathematical tool for reconstructing spatial coordinates is described in this section. This part of the dissertation is of compilation character and cites publications of other authors. The epipolar geometry is the basis of spatial reconstruction. We are able to obtain spatial information by using pairs of corresponding points. Corresponding points are image points which represent the same spatial point in two or more input images. The image of a particular spatial point is located at various positions in various images. These differences are given by mutual positions of the cameras. Actually, many various methods were proposed for executing each step of the reconstruction. However, describing all the existing approaches is not the aim of the section. One of the possible procedures for reconstruction is described in this section. The described methods are implemented to the created application (see

section 2.5) used for executing the experiments in my dissertation.

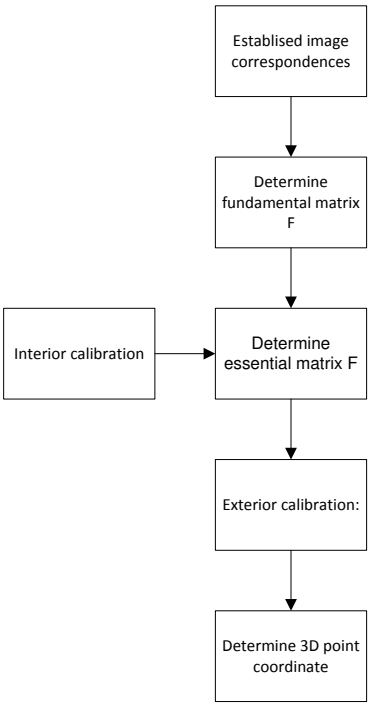


Fig. 2.1: The flowchart of the procedure for spatial coordinate reconstruction.

2.1.1 Procedure for model reconstruction

The process of reconstruction consists of a few fundamental steps. The reconstruction process can be described using the flowchart shown in Fig.2.1. The variant with interior calibration executed in advance is used in my practical implementation. More specifically, the interior calibration executed by using calibration patterns in our application. The reliable determination of he positions of the corresponding points $p_{1[x_1,y_1]}$ and $p_{2[x_2,y_2]}$ in both particular input images (I_1,I_2) is an essential assumption for accurate reconstruction of the spatial coordinates corresponding to the scene points *P[X,Y,Z]*. This process can be divided into two steps: finding significant points and matching the found significant points. A large part of my work is devoted to this topic. Chapters 2.2, 2.3 and 2.4, deals with this aspect. When points' correspondences are available, exterior calibration can be executed. Projective reconstruction can be obtained even without the knowledge of interior calibration. The last step is transformation to euclidean reconstruction (metric calibration). During this transformation, interior calibration or knowledge of the scene is used.

2.1.2 Interior calibration

Many methods were proposed for interior calibration. The offline method is used in the designed application (see. section 2.5). The application uses the open source toolbox Camera Calibration for MATLAB [83] for interior calibration. The toolbox utilizes the method proposed in [84]. In this chapter, this used method will be briefly described.

The used method does not have great demands on equipment and used a planar calibration template. It is necessary to capture a scene at least from three various camera positions. The template is often a chessboard with known properties. This approach has two phases. In the first step, the homography is determined by using Direct Linear Transformation(one of the first using DLT [85]. We consider only a linear relation without radial distortion in this step. The basis of the second step is optimizing estimated parameters of the camera. Optimization is based on the Maximum Likelihood criterion. The method assumes that the calibration template lies in the plane $Z = 0$ against reference coordinate. Therefore we can write [84]

$$\lambda \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \mathbf{K} \begin{bmatrix} r_1 & r_2 & r_3 & t \end{bmatrix} \begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix} = \mathbf{K} \begin{bmatrix} r_1 & r_2 & t \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix}. \tag{2.1}$$

Where r_i represents the individual columns of the rotation matrix $\mathbf{R}$. The symbol λ is an arbitrary unknown factor. The homography ($\mathbf{H}$) between calibration template and its image is determined up to this factor. The homography between the image of the point and its pattern can be written as

$$\lambda x = \mathbf{H}X, \tag{2.2}$$

$$\mathbf{H} = \mathbf{K} \begin{bmatrix} r_1 & r_2 & r_3 \end{bmatrix}. \tag{2.3}$$

The conditions for the interior parameters of the camera can be derived in the following way, because the columns of the rotation matrix are orthonormal.

$$\begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix} = \lambda \begin{bmatrix} \alpha_x & s & x_0 \\ 0 & \alpha^y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & h_{12} & t_1 \\ r_{21} & h_{22} & t_2 \\ r_{31} & h_{32} & t_3 \end{bmatrix}. \tag{2.4}$$

From condition $r_1^T r_2 = 0$, we can derive

$$h_1^T \mathbf{K}^{-T} \mathbf{K}^{-1} h_2 = 0, \tag{2.5}$$

where $\mathbf{K}^{-T}$ denotes transposition of the inverse matrix $\mathbf{K}^{-1}$ and from condition $r_1^T \cdot r_1 = r_2^T \cdot r_2$ we can write

$$h_1^T \mathbf{K}^{-T} \mathbf{K}^{-1} h_1 = h_2^T \mathbf{K}^{-T} \mathbf{K}^{-1} h_2. \tag{2.6}$$

Equations 2.5 and 2.6 represent two conditions which are asked by homography matrix $\mathbf{H}$ from the calibration matrix $\mathbf{K}$. In this method, the symmetric matrix is considered

$$\mathbf{B} = \mathbf{K}^{-T} \mathbf{K}^{-1} = \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{12} & B_{22} & B_{23} \\ B_{13} & B_{23} & B_{33} \end{bmatrix}. \tag{2.7}$$

When we designate columns of matrix $\mathbf{H}$ as h_j where j is the number of the row. subsequently, we can write

$$h_i^T \mathbf{B} h_j = v_{ij}^T b, \tag{2.8}$$

where

$$v_{ij} = [h_{i1}h_{j1}, h_{i1}h_{j2} + h_{i2}h_{j1}, h_{i2}h_{j2}, h_{i3}h_{j1} + h_{i1}h_{j3}, h_{i3}h_{j2} + h_{i2}h_{j3}, h_{i3}h_{j3}]^T. \tag{2.9}$$

From equation 2.5, we can derive relation

$$h_1^T \mathbf{B} h_2 = v_{12}^T b = 0, \tag{2.10}$$

and from 2.6

$$h_1^T \mathbf{B} h_1 - h_2^T \mathbf{B} h_2 = (v_{11} - v_{22})^T b = 0. \tag{2.11}$$

These two equations are valid for one homography defined by the homogeneous system of the equations

$$Vb = 0 \qquad \begin{bmatrix} v_{12}^T \\ (v_{11} - v_{22})^T \end{bmatrix} b = 0. \tag{2.12}$$

Five interior parameters of the camera can be determined by using n images and by constructing the system $Vb = 0$ where matrix V has the size 2n x 6. The minimal number of n is 3. The vector b is aligned to matrix B which specifies interior parameters as:

$$y_0 = \frac{B_{12}B_{13} - B_{11}B_{23}}{B_{11}B_{22} - B_{12}^2}, \tag{2.13}$$

$$\lambda = B_{33} - \frac{[B_{13}^2 + v_0\dot{(}B_{12}B_{13} - B_{11}B_{23})]}{B_{11}}, \tag{2.14}$$

$$a_x = \sqrt{\frac{\lambda}{B_{11}}}, \tag{2.15}$$

$$a_y = \sqrt{\frac{\lambda B_{11}}{(B_{11}B_{22} - B_{12}^2)}}, \tag{2.16}$$

$$s = \frac{-B_{12}\alpha^2\beta}{\lambda}, \tag{2.17}$$

$$x_0 = \frac{\gamma v_0}{\beta} - \frac{B_{13}\alpha^2}{\lambda}. \tag{2.18}$$

2.1.3 Exterior calibration

Two vectors are inputs of the exterior calibration. Each of the vectors contains position of the corresponding points in one of the input images. The outputs of the exterior calibration are rotation matrix $\mathbf{R}$ and translation vector $\mathbf{T}$. The mentioned parameters represent the mutual position of the two cameras. The projection matrix can be created using these parameters. The projective reconstruction of the scene can be obtained using the projective matrix. The most used algorithm for obtaining exterior calibration is 8- point algorithm [86]. The eight algorithms will be briefly described.

For a given set of images correspondences $\left(x_1^j, x_2^j\right), j = 1, 2...., n\,(n > 8)$, the 8-point algorithm recovers the rotation matrix and translation vector which satisfy

$$x_2^{jT}\hat{T}\mathbf{R}x_1^j = 0, j = 1, 2....., n. \tag{2.19}$$

We construct $A = [a^1, a^2, a^3, ..., a^n]^T \in \Re^{n\times 9}$ where a^j is gained from correspondences x_1^j and x_2^j as

$$a^j = x_1^j \otimes x_2^j \in \Re^9. \tag{2.20}$$

Then we compute singular value decomposition of A $[U, D, V] = SVD(A)$. We define $\mathbf{F}$ as the ninth column of matrix $\mathbf{V}$. Then, the vector is reshaped into a square 3 x 3 matrix . We have to execute the following operation in order to guarantee that fundamental matrix $\mathbf{F}$ has rank of matrix 2. The fundamental matrix is an algebraic expression of epipolar geometry. The fundamental matrix maps the position of the points in one image to its position in the second image. The fundamental matrix includes information about interior calibration. The essential matrix is a special form of the fundamental matrix. The essential matrix does not include information about interior calibration. We compute singular value decomposition of $\mathbf{F}$, $[U_1, D_1, V_1] = SVD(\mathbf{F})$. Subsequently, we set the smallest singular number in diagonal matrix $\mathbf{D}_1$ to zero. Then we obtain the required fundamental matrix $\mathbf{F}$ by composition

$$\bar{\mathbf{F}} = \mathbf{U}\bar{\mathbf{D}}\mathbf{V}^T. \tag{2.21}$$

Subsequently, we can determine matrix $\mathbf{R}$ and $\mathbf{T}$. In the first step we get the essential matrix $\mathbf{E}$ by using calibration matrix $\mathbf{K}$.

$$\mathbf{E} = \mathbf{K}'^{T}\mathbf{F}\mathbf{K}. \tag{2.22}$$

We need to find projective matrices of both camera $\mathbf{P_R}$ and $\mathbf{P_L}$. We can set one of the projective matricesto the beginning of coordinate system

$$\mathbf{P_L} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}. \tag{2.23}$$

Then, the projective matrix $\mathbf{P_R}$ contains $\mathbf{R}$ and $\mathbf{T}$. We can derive $\mathbf{P_R} = [\mathbf{RT}]$ from the essential matrix. $[U, S, V] = SVD(\mathbf{E})$ There are four different variants of the projective matrix (Fig. 2.2)

$$\mathbf{R} = UWV^T, \quad \mathbf{T} = U, \tag{2.24}$$

$$\mathbf{R} = UWV^T, \quad \mathbf{T} = -U, \tag{2.25}$$

$$\mathbf{R} = UW^TV^T, \quad \mathbf{T} = U, \tag{2.26}$$

$$\mathbf{R} = UW^TV^T, \quad \mathbf{T} = -U, \tag{2.27}$$

where $\mathbf{U}$ and $\mathbf{V}$ is obtained from $[\mathbf{U}, \mathbf{S}, \mathbf{V}] = SVD(\mathbf{E})$ and $\mathbf{W}$ is defined as

$$\mathbf{W} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{2.28}$$

The variants of matrix $\mathbf{P_R}$ can be geometrically represented (see Fig 2.2). The object is in front of both cameras only in one case. Consequently, matrix $\mathbf{P_R}$ which satisfies this condition will be selected.

2.1.4 Triangulation

When we have a set of corresponding points $\left(x_1^j, y_2^j\right)$, rotation matrix $\mathbf{R}$ and translation vector $\mathbf{T}$, then we can reconstruct the spatial coordinates of the image points. Many methods can be found for executing triangulation in literature. A well written summary and comparison of these methods can be found in [87]. I used the linear triangulation method in my application. The 3D structure X for each j=1,2,3....,n can be estimated as follows. Denote the individual rows in the projective matrix $\mathbf{P_R}$ as p_r^1, p_r^2,p_r^2 and $\mathbf{P_L}$ as p_l^1, p_l^2,p_l^2 then

$$\mathbf{A} = \begin{bmatrix} x_1^j p_l^{3T} - p_l^{1T} \\ y_1^j p_l^{3T} - p_l^{2T} \\ x_2^j p_l^{3T} - p_r^{1T} \\ y_2^j p_l^{3T} - p_r^{2T} \end{bmatrix}. \tag{2.29}$$

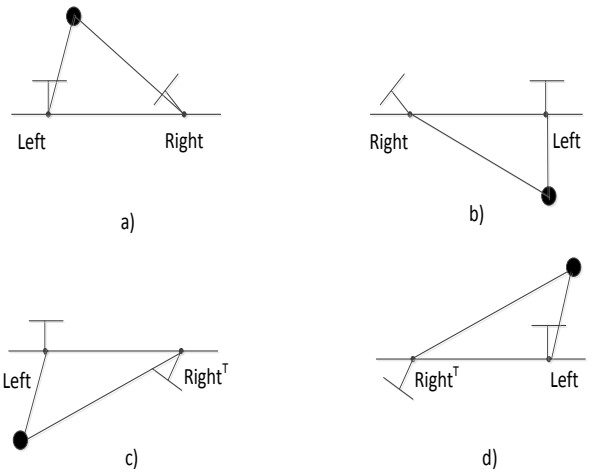


Fig. 2.2: The geometric representation of the various variants of the projective matrix.

These equations define X only up to an undetermined scale factor λ.Subsequently, the projective structure can be recovered as the least-squares solution of a linear system of equations $\mathbf{A} \cdot X_p = 0$. The system can be solved using singular value decomposition $[ua, sa, va] = svd(\mathbf{A})$. Then, the space coordinates of the points are obtained from the last column of va. Then, we normalized fourth coordinate of X_p^j to 1. The unknown scales λ_1^j are the third coordinate of the homogeneous representation of X_p^j.

2.2 Comparison of commonly used methods for finding corresponding points

The fundamental step in the process of model reconstruction is finding the corresponding points. We can divide this procedure into two steps. Firstly, we need to find significant points in both images. Subsequently, we have to determine which points represent the same point in the scene (matching). Finding significant points is executed by detectors. A significant point is a point which can be found repeatedly. The detectors have to comply with some conditions. The detector has to be invariant to

- translation,
- rotation,
- change of the scale,
- change of the intensity and contrast,
- change of the view angle.

The purpose of comparing commonly used methods for finding corresponding points is to select one of them which will be further used in the test. The most frequently used methods were tested. The Harris detector is a detector among other methods which can be assigned as descriptors. If we use a Harris detector, an appropriate method has to be used for finding corresponding pairs of points in both images. Performance of the method is compared using the reliability of the finding corresponding points.

2.2.1 Harris detector

The Harris detector is frequently used nowadays, although it was first published by Chris Harris in 1988 [31]. The fundamental idea is finding a the place in image, in which gradient is changing in two directions. Therefore, the Harris detector is rotation invariant. It means that rotation of the image does not have any influence on finding significant points. The Harris detector can be denoted as the successor of the Moravec detector. Calculation of the gradient can be disturbed by noise. For noise elimination, the Harris detector uses window in Gauss function. The Gauss function is described in and given by equation 2.30

$$g\left(x,y\right) = exp^{\left(\frac{\left(x^2+y^2\right)}{2\sigma^2}\right)}, \tag{2.30}$$

where σ is standard deviation, which specifies the smoothness of the image. Therefore, the Harris detector local autocorrelation function E(x, y) is [31]

$$E\left(x,y\right) = \sum_{W}\left[I(x_i,y_i) - I(x_i+\Delta x, y_i+\Delta y)\right]^2. \tag{2.31}$$

Where Δx a Δ are elementary shifts, *I(x,y)* denotes image function, W indicates the window in which a significant point is found. Further points (x_i, y_i) are points in this window with center at (x,y). Subsequently, shift of the image function is approximate by the first two members of the Taylor's series [31]

$$f\left(x+u,y+v\right) \approx f\left(x,y\right) + \left[f'_x(x,y)f'_y(x,y)\right]\begin{bmatrix}\triangle x\\ \triangle y\end{bmatrix}. \tag{2.32}$$

Where f_x' and f_y' are partial derivation v x and y. We substitute 2.32 into 2.31. Mathematical operations are executed. Subsequently, the following equation is obtained [31]

$$E(x,y) = \sum_{W}\left(\begin{bmatrix}\frac{\partial I(x_i,y_i)}{\partial x} & \frac{\partial f(x_i,y_i)}{\partial x}\end{bmatrix}\begin{bmatrix}\Delta x\\ \Delta y\end{bmatrix}\right)^2, \tag{2.33}$$

after the next operation, we obtain the following equation [31]

$$E(x,y) = [\Delta x \Delta y]\cdot\begin{bmatrix}\sum_W \frac{\partial^2 f(x,y)}{\partial x^2} & \sum_W \frac{\partial^2 f(x,y)}{\partial x \partial y}\\ \sum_W \frac{\partial^2 f(x,y)}{\partial x \partial y} & \sum_W \frac{\partial^2 f(x,y)}{\partial y^2}\end{bmatrix}\cdot\begin{bmatrix}\Delta x\\ \Delta y\end{bmatrix}. \tag{2.34}$$

this can be rewritten as [31]

$$E(x,y) = [\Delta x \Delta y]\, Q(x,y) \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix}. \tag{2.35}$$

From equation 2.35, the autocorrelation matrix $\mathbf{Q}$(x,y) is determined. The matrix is calculated by using partial derivation. For clarity and simplicity, matrix $\mathbf{Q}$*(x,y)* can be rewritten in the following form [31]

$$\mathbf{Q}(x,y) = \begin{bmatrix} \sum_W \frac{\partial^2 f(x,y)}{\partial x^2} & \sum_W \frac{\partial^2 f(x,y)}{\partial x \partial y} \\ \sum_W \frac{\partial^2 f(x,y)}{\partial x \partial y} & \sum_W \frac{\partial^2 f(x,y)}{\partial y^2} \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix}. \tag{2.36}$$

The corner point can be found by using this matrix. After calculating the matrix, response function R is calculated by the following relation [31]

$$R = \det \mathbf{Q}(x,y) - \kappa \cdot trace^2(\mathbf{Q}(x,y)), \tag{2.37}$$

where κ is constant. The best value of this constant was experimentally determined in the range 0.04-0.06. The matrix determinant (det($\mathbf{Q}$)) and matrix trace (trace($\mathbf{Q}$)) is determined using eigenvalues of the matrix $\mathbf{Q}(x,y)$ [31]

$$det(\mathbf{Q}(x,y)) = \lambda_1 \lambda_2 = AC - B^2, \tag{2.38}$$

$$trace(\mathbf{Q}(x,y)) = \lambda_1 + \lambda_2 = A + C. \tag{2.39}$$

Subsequently, the response function can be expressed as [31]

$$R = (A \cdot C - B^2) - \kappa (A + C)^2. \tag{2.40}$$

The local extreme of the response function R are denoted as significant points. The decision is executed depending on whether $R(x,y)$ exceeds the selected threshhold T.

2.2.2 Scale-invariant feature transform

This method was firstly published by D.G. Low in 2004 [30]. This corner detector is concurrently a descriptor.Therefore, the algorithm describes found points by some features (descriptor). Each point is described by this descriptor. The descriptor vector consists of 128 integer numbers. The big advantages of SIFT is obvious from its name. SIFT is invariant to changes of scale. The algorithm is further invariant to translation, rotation, affine deformation and partly brightness transformation. Matching points is executed by comparing the descriptors. Execution of the method can be divided into the following steps:

1. detection of candidates on being the significant points in scale space,

2. elimination of unstable candidates,
3. determination of the orientation of each point,
4. generation and assignment of the descriptor to each point.

Subsequently, a brief description of each step follows. Detection of the candidates is executed in scale space. The process is illustrated in Fig. 2.3. The used approach ensures invariance to the change of scale. Practically, scale space is obtained by executing detection in a few various resolutions of the input image. We apply the LoG (Laplacian of Gaussian) filter on the input image. Possibly, DoG (difference of Gaussian) can be used. The filtration is executed by convoluting the input image with the Gaussian filter. Filtration is repeated for the same image (same resolution) with various standard deviation σ. Subsequently, differences of the images acquired by filtering with various σ are calculated. Local extreme in $\mathbf{D}(x, y, \sigma)$ are denoted as candidates on the significant points. For finding a local extreme, the value of the pixel is compare with the pixel in its neighborhood in all scales. A large number of the candidates is obtained.

Subsequently, the elimination of unstable points is executed. Points which lay along the edge are eliminated. For this purpose, the Hessian matrix is used. The Hessian matrix contains the second derivation of the image. We have to determine the threshold. Subsequently, we make the decision whether the points are regular or if they lay on the edge (unstable). Points with insufficient contrast are eliminated too.

In the next step, orientation is assigned to the significant points. This process is based on using the orientation of the gradient in the point's neighborhood. For an examined point, a histogram of the orientation is built. The histogram has 36 bins ensuring a coverage of 360 degrees. The dominating orientation is determined as the peak of the histogram. The orientation of the point serves to ensure invariance to rotation and is represented by the orientation of a few of the most significant gradients in the point's neighborhood.

Further, descriptors for each point are calculated. In the first step, the neighborhood of the particular significant point is divided to an n x n square. The histogram of the orientation of the gradient is construed for each square area. In algorithm SIFT, n is equal to 4. Therefore, we have 16 areas, each of them is described by a histogram (8 bits). The subsequently size of the descriptor is 16 x 8 = 128.

Finally, we have a set of the significant points in both input images. Significant points are described by a descriptor and we assign a corresponding point using this descriptor. Points with the most similar descriptor are denoted as corresponding points. The comparison of two pints is executed by comparing their descriptor. The comparison is performed by calculating the Euclid distance. Unfortunately, small speed is disadvantages of the SIFT. This algorithm is not used in the application in

real time.

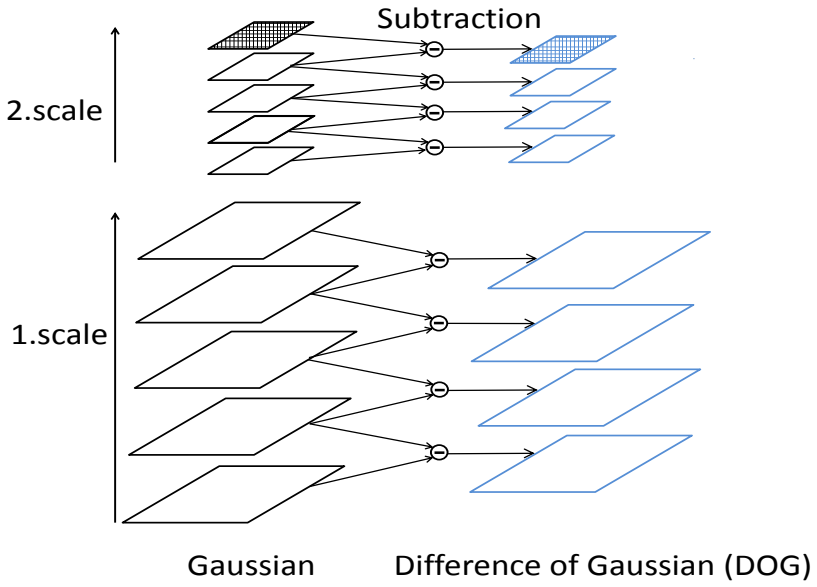


Fig. 2.3: Basic principal of SIFT: change of scale and blurring.

2.2.3 Speeded up robust feature

Speeded–Up Robust Features (SURF) was introduced in 2006 [29]. This method was inspired by SIFT. An effort to accelerate of the process is the reason for developing new methods on the same base. Acceleration is achieved by using the approximation of the Hessian matrix. Using this approximation leads to using an integral image [88], it decreases computing difficulty. SURF uses a smaller descriptor. This is another thing which increases speed. The integral image is a simple structure for quickly finding the sum value in an arbitrary rectangular area in images. The integral image has the same size as the original image. The calculation is executed by a function which ensures that the sum of the pixel in the area can be determined based on the value of the points which are around this area.

2.2.4 Experiment and results

In this section, the executed test and its results are described. We tested algorithms for finding significant points and subsequently assigning corresponding points in both images. In the test, images from the Middlebury stereo dataset were used [89]. The database contains 21 images. Miniatures of the 20 images from database are shown in Fig. 2.4. We used these images because the database contains even their true depth map. The true depth maps are important in evaluating the correctness of the determined correspondences. The images, even depth maps, have resolution

1310 x 1112. The calculation of the objective parameters of the image was included in the test. We investigated the impact of the following properties of the image to the correctness of the correspondences:

- structural Similarity Index Measure (SSIM) [90],
- spatial Activity (SA),
- frequency Activity (FA),
- correlation coefficient(CC),
- standard Deviation (SD),
- EDGE,
- local Entropy (LE),
- local Range (LR),
- contrast (CO).

These parameters are briefly described. The SSIM index is a method for measuring the similarity between two images. The SSIM index is commonly used in full reference metrics with a reference image for evaluating image quality. We used the SSIM index for measuring the similarity between the left and right image. The SSIM index is calculated using the following equation [90]

$$\mathrm{SSIM}\left(x,y\right)=\frac{\left(2\mu_x\mu_y+c_1\right)\left(2\delta_{xy}+c_2\right)}{\left(\mu_x^2+\mu_y^2+c_1\right)\left(\delta_x^2+\delta_y^2+c_2\right)}, \tag{2.41}$$

where μ is average value, η is standard deviation (variance), c_1 and c_2 are constants.

The spatial activity gives information about the frequency of the changes of intensity. SA is calculated as the mean change between the adjoining pixels in the vertical and even horizontal direction.

$$\mathrm{SA}=\frac{\sum_{i=0}^{m}\sum_{j=0}^{n}\left[\left(I\left(x_{i-1},y_j\right)-I\left(x_i,y_j\right)\right)+\left(I\left(x_i,y_{j-1}\right)-I\left(x_i,y_j\right)\right)\right]}{mn}, \tag{2.42}$$

where m and n are dimensions of the image, x_j and y_j are particular position in the image and $I\left(x_i,y_j\right)$ is the value of the pixel at a particular position.

The frequency activity brings information about the presence of higher harmonics. Higher harmonics informs us about edges. Firstly, frequency representation of the image is obtained by using Fourier transformation. In the next step, the higher harmonics are filtered. Subsequently, the ratio of higher harmonics and all harmonics are calculated. The correlation is the correlation coefficient between the left and right image. Therefore, CC gives information about the relationship between the left and right image. The correlation coefficient is calculated by using the following equation [91]

$$\mathrm{CC}=\frac{1}{n-1}\sum_{i=0}^{n}\left(\frac{X_i-X}{s_x}\right)\left(\frac{Y_i-Y}{s_y}\right), \tag{2.43}$$

where X and Y are means of the image intensity

$$X = \frac{1}{n}\sum_{i=0}^{n} X_i Y = \frac{1}{n}\sum_{i=0}^{n} Y_i, \tag{2.44}$$

and x_x, s_y are the standard deviation of the images.

$$s_x = \sqrt{\frac{1}{n-1}\sum_{i=0}^{n}(X_i - X)^2}, s_y = \sqrt{\frac{1}{n-1}\sum_{i=0}^{n}(Y_i - Y)^2}. \tag{2.45}$$

The term EDGE indicates the number of edges in an image. The parameter gives information: how many pixels were marked as edge. The Canny edge detector is used [92] for this purpose.

The entropy generally represents the degree of uncertainty of a system (image in this case). The local entropy (LE) [93] is determined as the mean value of the entropies calculated separately for all pixels in the image [94].

The local range (LR) represents the local dynamic in the image. The range of the individual pixel is given by the difference of minimal and maximal value in its neighborhood. Subsequently, the parameter representing the whole image is determined as the average value of the ranges calculated separately for all pixels.

The pairs of the parameters CA and SSIM, FA and EDGE deal with similar information but in various forms. This fact was the aim. We wanted to test different representations of same the properties. Subsequently, the evaluation of the relation between objective parameters and performance of the algorithm was executed. The process of the test can be described by the following steps:

1. Assignment of the set of corresponding points.
2. Calculation of the horizontal disparity for each pair of corresponding points.
3. Comparison of the horizontal disparity with the value of the appropriate pixel in the true disparity map.
4. Calculation of the ratio of incorrect and correct correspondences.
5. Calculation of the objective parameters of the images.
6. Final evaluation of the obtained data.

Assignment is executed by using tested algorithms (Harris detector, SIFT, SURF). Inputs are left image i_{left} and right image i_{right}. The outputs are a set of significant points in the form of two vectors called Pos_{left} and Pos_{right} (one for each image). The calculation of the horizontal disparity is calculated as the difference between horizontal coordinates (rows) of the corresponding points

$$disparity_{found}\,(i)) = y_{1,i} - y_{2,i}, \tag{2.46}$$

where i is the order number of the corresponding points, y_1 represents horizontal position (column) in the first image and y_2 represents horizontal position in the

second image. The comparison of the true disparity ($disparity_{true}$) given by the true disparity map and $disparity_{found}$ is ultimately calculated by using

$$difference_{disparity}\,(i)) = |disparity_{true} - disparity_{found}|. \tag{2.47}$$

The appropriate correspondence is denoted as incorrect if the $difference_{disparity}$ exceeds the threshold. The threshold was experimentally determined to equal to 5.

The main aim is mutual comparison of algorithms. Fig. 2.5 shows the reliability of determining correspondences by different algorithms for individual images. The reliability in percentage is obtained by the following formula

$$Reliability = \left(1 - \frac{WrongCorrespondences\_number}{FoundCorrespondences\_number}\right) 100. \tag{2.48}$$

Obviously, SIFT provided the best results. This fact is confirmed when the average success rate is calculated. These average values and standard deviations are in Tab. 2.1. The best results are provided by SIFT with an average success rate of 97.43 %. On the contrary, the worst results are provided by the algorithm Harris detector with an average success rate of 80.72%. Tab. 2.2 contains objective parameters of all images used in the experiment. The values were rounded to 4 significant digits. All parameters are dimensionless numbers. We reveal that reliability is dependent on selected parameters. However, the dependency on the individual parameters is weak. Therefore, the determination of the strongest indicator for predicting the reliability of finding corresponding points was the aim in the next step. The experiment revealed that some of these parameters have an impact on the success rate of finding corresponding points. The following parameters belong to this group: spatial activity, frequency activity, local range, local entropy, the number of edges. On the contrary, other parameters do not have an impact on finding corresponding points. The level of significance was improved by combining of relevant parameters. We designed the parameter K_r which is given by the following relation

$$K_r = \frac{\mathrm{SA} \cdot \mathrm{LE} \cdot \mathrm{Edge}^{0.25}}{\mathrm{FA} \cdot \mathrm{LR}}. \tag{2.49}$$

Parameter K_r serves to describe the images. We can estimate the probability of good reliability of finding corresponding points. When the value of K_r increases, then the probability of good reliability also increases. This fact is obvious from Fig.2.6-Fig.2.8, where parameter K is on the horizontal axis and the reliability of finding corresponding points is on the vertical axis. Parameter K_r is normalized to a range from 0 to 1 for a better illustrative nature. The real range of K_r, for the used images, is from 8.46 to 49.79. This relation was obtained experimentally. At first, we investigated the influence of more image properties than was mentioned above.

However, we discovered that some properties have no impact on the reliability. The parameters which have significant impact were investigated further. Subsequently, we evaluated if reliability is directly or indirectly proportional to a particular parameter. The parameters directly proportional to reliability were placed to the numerator. On the contrary, the parameters indirectly proportional were placed to the denominator. The exponents for individual parameters were determined according to the degree of dependency. The reliability is indirectly dependent to the LR; this fact is surprising. In the last step, the ideal weight of the individual parameter for strong dependency was found. The dependencies of reliability of finding corresponding points by the used method on the parameter K_r is shown in figures 2.6, 2.7 and 2.8. Obviously, the reliability increases with increasing parameter K_r for every method.

Method	**HARRIS**	**SURF**	**SIFT**
Average success rate [%]	80.72	82.80	97.43
Standard deviation [%]	12.57	18.18	4.11

Tab. 2.1: Comparison of reliability of finding corresponding points by commonly used methods SURF, SIFT and Harris detector for the used set of images 2.4 [89].

no.	**SSIM**	**SA**	**EDGE**	**FA**	**CC**	**SD**	**LE**	**LR**	**CO**
1	0.9971	0.026	73242	0.0219	0.4751	0.1615	4.5902	0.0928	0.9176
2	0.999	0.0123	66095	0.0226	0.579	0.1092	3.1621	0.0419	0.7696
3	0.9972	0.0116	24208	0.0228	0.7427	0.2045	3.0963	0.0396	0.9584
4	0.997	0.0106	18597	0.0239	0.6487	0.1833	3.3996	0.0369	0.8874
5	0.9973	0.0078	13713	0.0269	0.6102	0.1672	2.7855	0.0263	0.9075
6	0.9929	0.0112	19158	0.0225	0.2158	0.1905	3.3035	0.039	0.9788
7	0.9979	0.0255	224843	0.0254	0.1262	0.1119	4.711	0.0933	0.6586
8	0.9943	0.0157	27221	0.0264	0.2104	0.1813	4.1136	0.0567	0.9469
9	0.995	0.0206	44018	0.0263	0.5443	0.2076	4.633	0.074	0.9605
10	0.9951	0.0204	104341	0.023	0.3252	0.1663	4.333	0.0741	0.9146
11	0.9936	0.0057	9676	0.0319	0.0779	0.1537	2.8447	0.0203	0.8667
12	0.998	0.0057	21051	0.0257	0.5464	0.13	2.3331	0.0193	0.8879
13	0.9986	0.0054	14301	0.0262	0.5762	0.1136	2.3456	0.0179	0.857
14	0.9959	0.0124	25475	0.0204	0.671	0.223	2.6986	0.0417	0.9667
15	0.9959	0.0126	24491	0.0207	0.646	0.2158	2.6795	0	0.9563
16	0.9967	0.0124	42474	0.0207	0.3446	0.1453	2.6074	0.0433	0.9057
17	0.9947	0.0043	13273	0.0308	0.5766	0.2218	1.8879	0.0152	0.8075
18	0.9997	0.0185	25818	0.0236	0.4133	0.158	4.4504	0.0655	0.8631
19	0.9957	0.019	28117	0.0234	0.3591	0.1623	4.5096	0.0673	0.9503
20	0.9985	0.0116	11831	0.0215	0.5088	0.108	3.8031	0.0393	0.7211
21	0.9994	0.0089	11444	0.0261	0.8425	0.1338	3.3517	0.0311	0.7874

Tab. 2.2: Objective parameters of the images from the used set of images 2.4 [89].

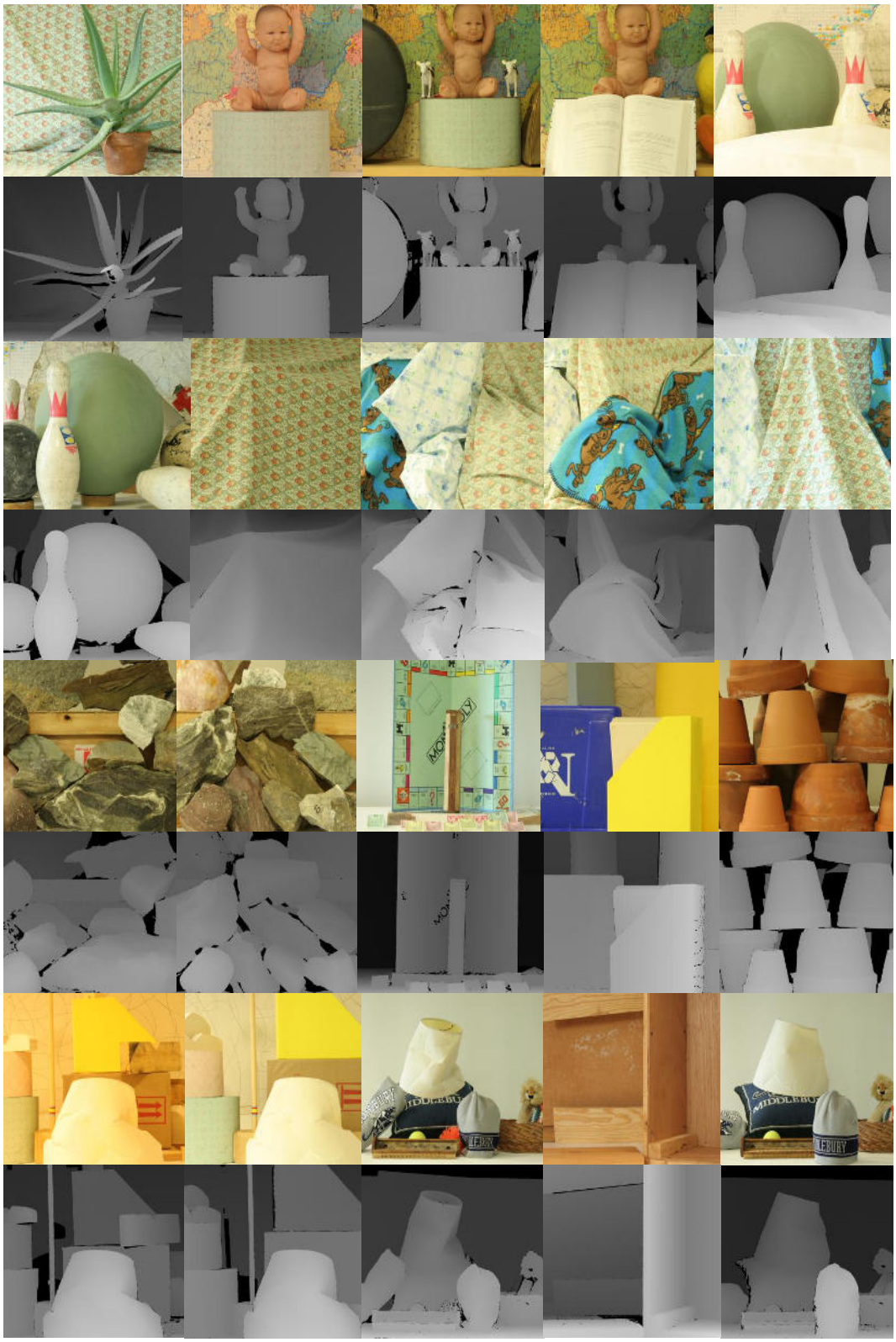


Fig. 2.4: Miniature of the images used in the experiment and their depth maps [89].

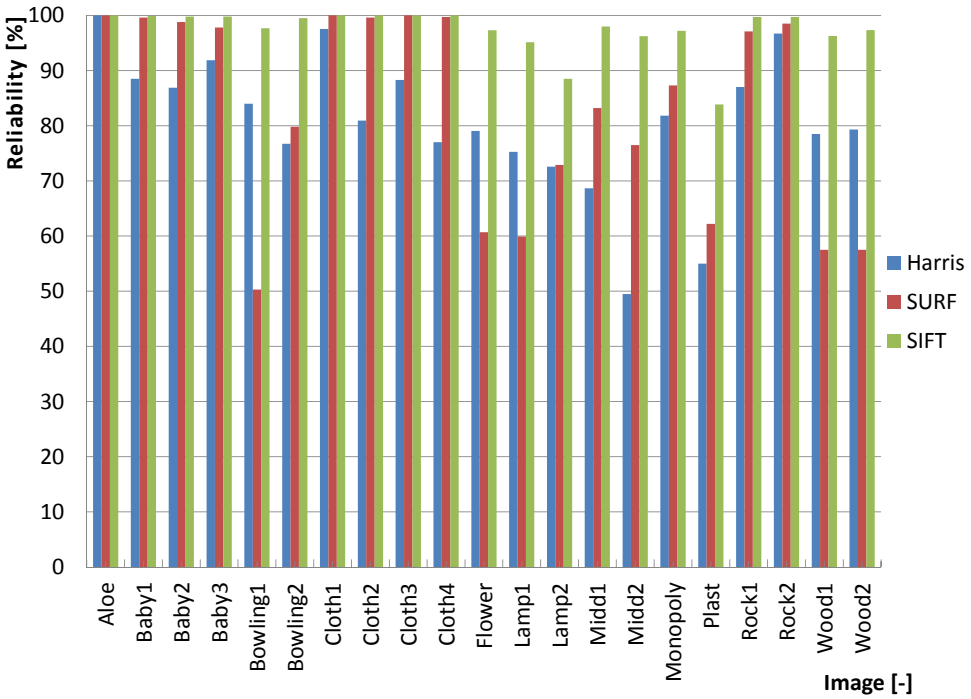


Fig. 2.5: The reliability of finding corresponding points by algorithms SURF, SIFT and Harris detector for an individual image from the used database [89] (see Fig. 2.4).

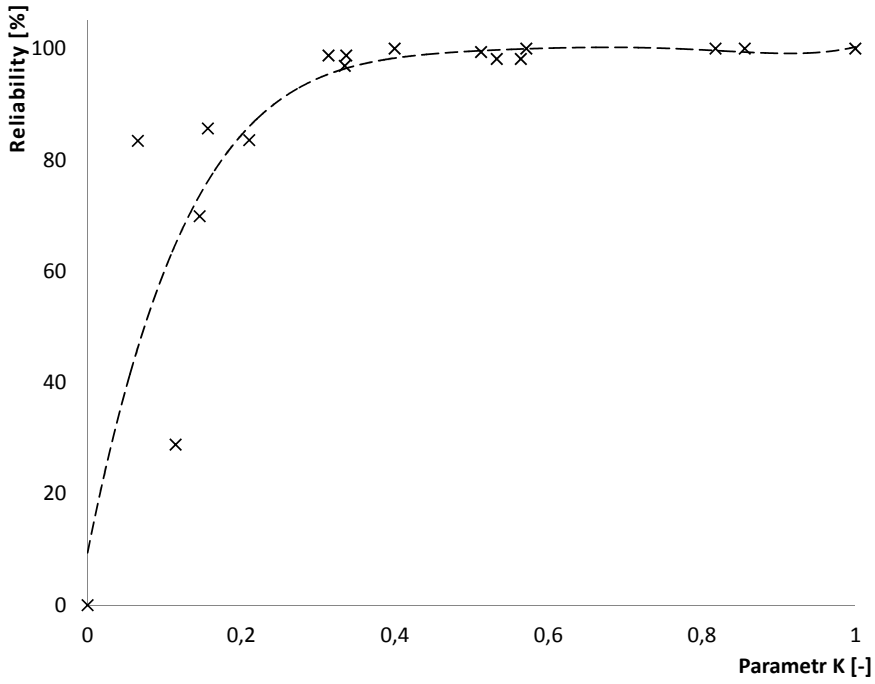


Fig. 2.6: The dependency of the reliability of finding corresponding points by the SIFT detector on the parameter K_r for individual images from the used database [89] (see Fig. 2.4).

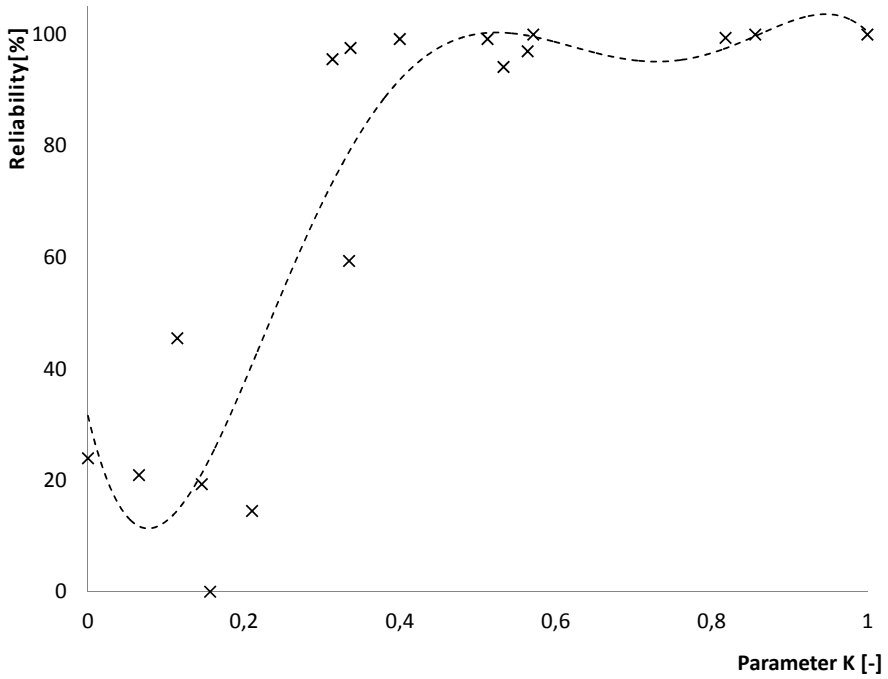


Fig. 2.7: The dependency of the reliability of finding corresponding points by the SURF detector on the parameter K_r for individual images from the used database [89] (see Fig. 2.4).

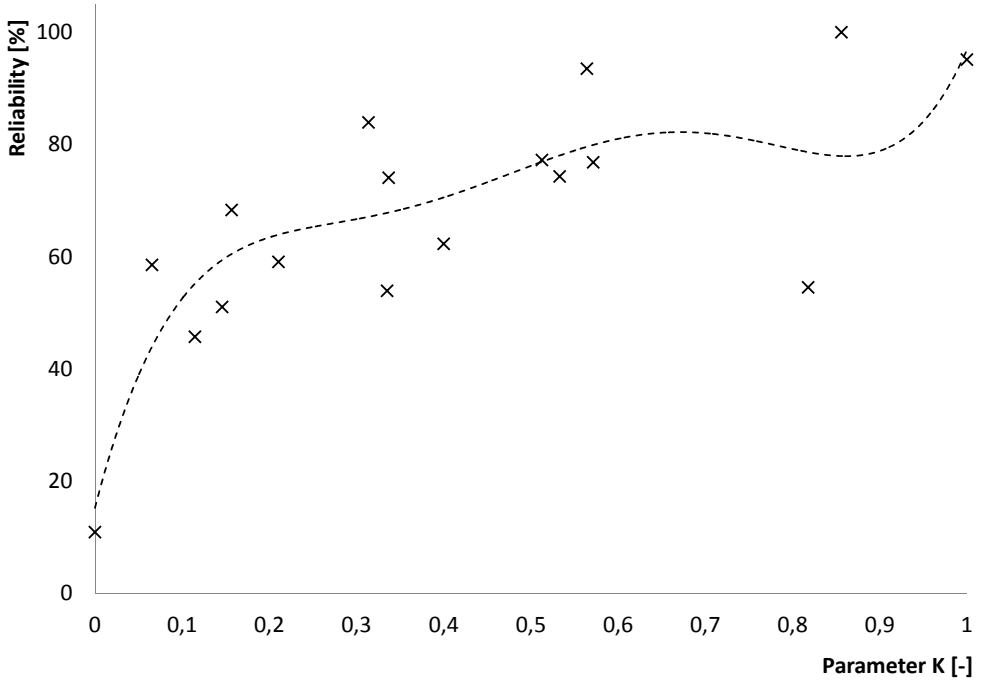


Fig. 2.8: The dependency of the reliability of finding corresponding points by the Harris detector on the parameter K_r for individual images from the used database [89] (see Fig. 2.4).

2.3 Proposed new method for correspondence of the selected point

2.3.1 Fundamental idea

Finding the corresponding point $p_{2,sel}\,(x_{2,sel}, y_{2,sel})$ in the right (second) image I_2 for selected point $p_{1,sel}\,(x_{1,sel}, y_{1,sel})$ in the left (first) image I_1 presents the main actual problem. The methods described in chapter 2.2 are used for finding corresponding points. However, these algorithms found matching only for significant points. Therefore, the user cannot determine for which points correspondences are found. A large area without correspondences can arise. Methods based on comparing the similarity of a point's neighborhood can be used for finding correspondences of a specific point. Usable methods are, for example, SAD, SAS, correlation and mutual information. Unfortunately, these methods can fail (find incorrect correspondences) in areas mentioned above (regular texture, without contrast). In this chapter, I will propose an algorithm, which will solve this task. Subsequently, the proposed algorithm is tested. The algorithm can be used for finding correspondence for a certain selected point or for thickening the net of significant points.

The inputs of the proposed process are image coordinates $x_{1,sel}, y_{1,sel}$ of the selected point in the left image I_1. The output are image coordinates $x_{2,sel}, y_{2,sel}$ of the corresponding point in the image I_2. The proposed procedure is based on probability. The basic principle of the method is using the following hypothesis.**If selected point P_{sel} is located in image area** A_{im}, which has certain depth $depth_{A_{im}}$, then there exists a high probability that point P_{sel} has equal depth $depth_{P_{sel}} = depth_{A_{im}}$. We will specify this fundamental assumption consecutively. We do not consider a uniform depth of the area. The depth of the area is defined using the depth of a few points belonging to the area in which the selected point is located. The depth of the neighborhood is represented by the depth of a few discrete points with small distance from the selected point. The hypothesis can be adjusted to the following form. **If we reliably know the depth of a few points in the selected point P_{sel} neighborhood, then we can determine the depth of the selected point with certain reliability**. The reliability is affected by a few factors. The determination of the depth is a difficult task in general conditions. Therefore, the hypothesis will be adjusted. We want to achieve the assumption which will be used only information directly obtained from images. The horizontal parallax and vertical parallax (further, summarizing label parallax will be used) are such information. The parallax gives information about the change of position of a point between both images. The concept **movement of the image point** (m_x, m_y) can be defined if we know image coordinates of the corresponding point in the images. The movement

of the point represents the change of the position of the image of the spatial point P in two various images of the scene. The movement of the point is determined as the difference of the appropriate image coordinates

$$m_x = x_{1,sel} - x_{2,sel}, \tag{2.50}$$

$$m_y = y_{1,sel} - y_{2,sel}. \tag{2.51}$$

Consequently, we can write:

If we know parallax of a few points in the neighborhood of the selected point then we can determinate parallax of the selected point with certain probability without knowledge of its corresponding point.

Therefore, calculating image coordinates of the selected point in the right image is possible. In consequence, the corresponding point for the selected point can be found if we know a sufficient number of point correspondences in its neighborhood.

Subsequently, the algorithm for practical implementation of this hypothesis was proposed. Its description is in section 2.3.2. Some conditions and assumptions are required. The basic assumption is knowledge of the reliable correspondences in the neighborhood of selected point. The required point correspondences can be found by various methods. In our practical implementation, the algorithm SURF was used.

2.3.2 Practical implementation

The flowchart of the algorithm is shown in Fig. 2.9. The inputs are image coordinates of the selected point $x_{1,sel}, y_{1,sel}$ in the left image. In the first step, the significant points are searched in a restricted neighborhood of the selected points in left and right images. The correspondences are found only in the restricted area of the image, respect the fact that only correspondence close to the selected point are important in the proposed algorithm. Subsequently, the Euclid distance is determined between individual significant points and a selected point in the left image. Consequently, a certain number of the closest significant points is selected. The number of used points has significant influence on the achieved results. If we use more points, then the probability increases that some of the selected points will lie outside the area with the same depth. The described situation should affect the results in a negative way. On the contrary, the probability of error exists if fewer points are used. The choice of five points was proved as a good compromise during experiments. The results of this step creates a set of significant points in the neighborhood of the selected point in the left $p_{1,SURF}$ and even in right image $p_{2,SURF}$.

In the next part of the procedure, we make the decision whether it is necessary to supplement the set of points with extra points. The decision is made by a trained

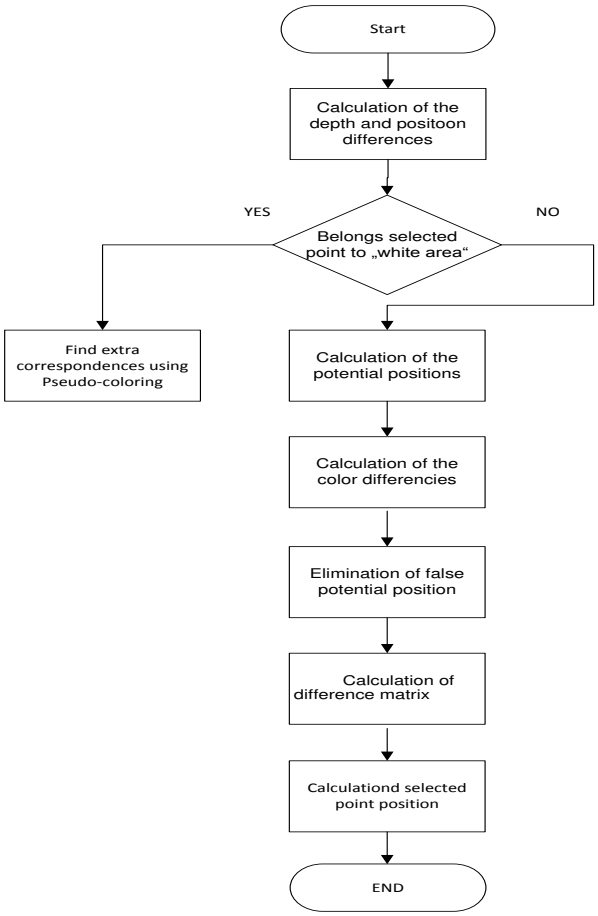


Fig. 2.9: The flowchart of the proposed system for finding a corresponding point for a selected point.

artificial neural network whose inputs are the depths of near feature points and their distances from the selected point. In the instance that the point lays in a dangerous area (too few correspondences found by SURF), adding extra information is necessary for obtaining accurate results. During the first test, the information is added using manual determination of auxiliary correspondences. However, this approach is unusable in practice. The next possibility is finding corresponding points by another method than SURF, which may found correspondences in the area. The last approach is finding corresponding points in another color model than true RGB (for example in HSV). In the next phase of research, we decided to use conversion to pseudo- color for finding new correspondences (see section 2.4).

In the case that a point does not belong to a danger area, then the algorithm continues by calculating the potential position of the selected point in the right image. Potential positions are the input to the last step of the algorithm for determining corresponding points. The second input is the position of the selected point in the left image. The calculation of the potential positions is based on the knowledge

of M_{SURF} and the position of the selected point in the left image $x_{1,sel}, y_{1,sel}$. The movement M_{SURF} represents the change of the position of the significant points between left and right images. For calculating M_{SURF} the following equation is used

$$M_{SURF} = p_{1,SURF} - p_{2,SURF}. \tag{2.52}$$

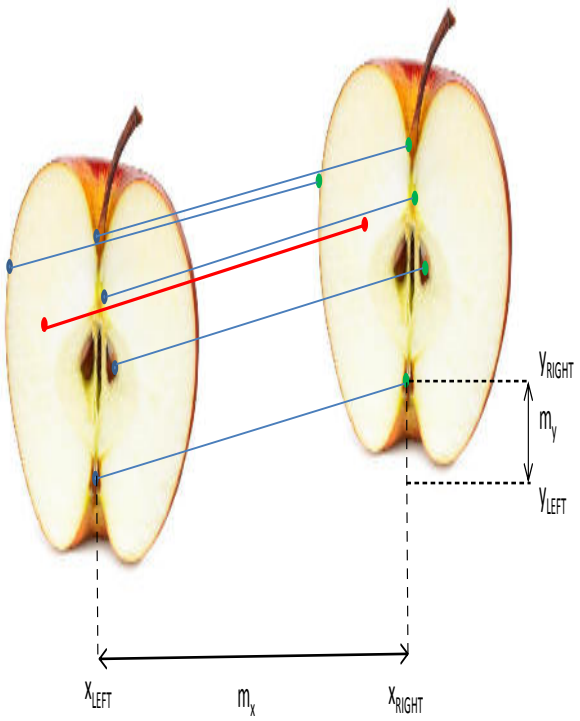


Fig. 2.10: Schematic drawing of finding the potential position of a selected point in the right image.

Then, potential positions $p_{2,potential}(x_{2,potential}, y_{2,potential})$ are calculated by using following equation, which represents the implementation of the basic hypothesis. The situation is schematic illustrate in Fig.2.10. The practical situation is shown in Fig.2.11

$$p_{2,potential}(x_{2,potential}, y_{2,potential}) = p_{1,sel}\,(x_{1,sel}, y_{1,sel}) - M_{SURF}. \tag{2.53}$$

Subsequently, we calculate the difference in color of the selected point in I_1 and of its potential position in I_2 ($diff_{color}$) using equation 2.54. The aim of this operation is identifying the unreal (wrong) potential position.

$$diff_{color} = \sqrt{(R_{sel} - R_j)^2 + (G_{sel} - G_j)^2 + (B_{sel} - B_j)^2}, \tag{2.54}$$

where $diff_{color}$ is the resulted difference of the color components, where $R_{sel}, G_{sel}, B_{sel}$ represent color components of the selected pixel P_{sel} in the left image and R_j, G_j, B_j

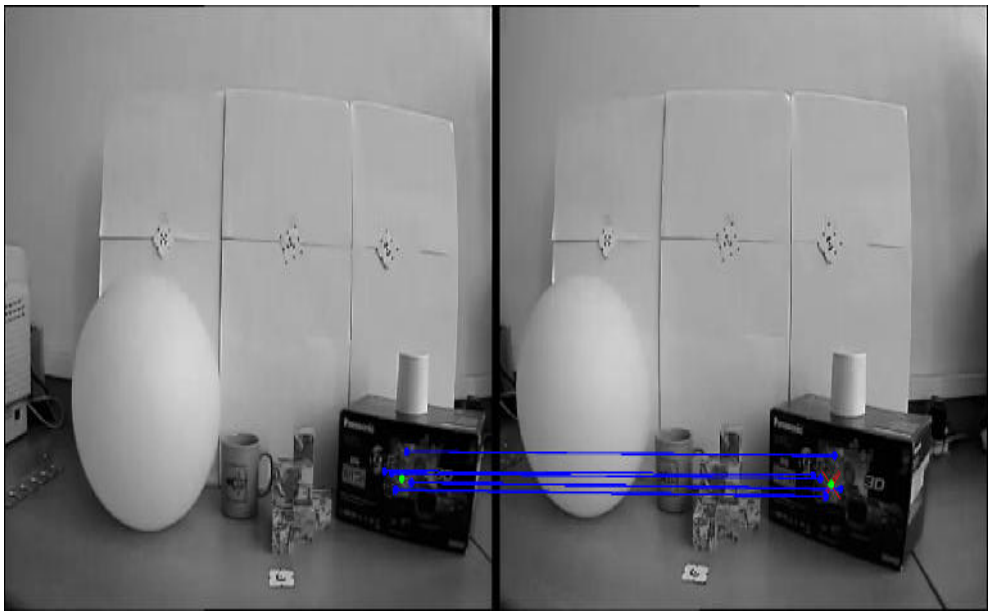


Fig. 2.11: Finding the position of the selected point in the right image.

represent color components of the pixels lying on potential positions of the corresponding point in the right image.

We eliminate the points whose difference exceeds a predefined threshold. This rule eliminates all potential pixels whose color is not sufficiently similar to the color of the selected pixel (in the left image). In this case, only one potential position remains in the right image after this operation. We decide that just this position is correct and we can calculate the spatial coordinates of the selected point. If more than one point remains, we continue with the next steps to obtain a reliable corresponding point. In the next step we calculate differences between individual potential positions ($dif f_{i,j}$) using the following equation

$$dif f_{i,j} = \sqrt{(x_i - x_j)^2 + (y_i - i_j)^2}, \tag{2.55}$$

where i, j are indices of the potential positions and x, y are image coordinates. Consequently, we obtain the $\mathbf{Difference_{matrix}}$ which contains individual differences:

$$\begin{vmatrix} dif_{1,1} & dif_{1,2} & ... & dif_{1,j} \\ dif_{2,1} & ... & ... & ... \\ ... & ... & ... & ... \\ dif_{i,1} & ... & ... & dif_{i,j} \end{vmatrix}. \tag{2.56}$$

The following step works with the $\mathbf{Difference_{matrix}}$, we can determine the layout of the points. An illustrative image in Fig. 2.12 shows different possible situations.

Depending on the situation, we calculate the final position of the selected point in the right image. We can define three basic situations:

- two points remain: the final position is given by the weighted average of their positions,
- two close points and one farther point remain: firstly, average position from nearby points is calculated, then weighted average with the farther point is calculated(farther point has a smaller weight),
- two pairs of points remain: firstly, the average position from the nearby points is calculated. Then, the weighted average from the averaged positions of both pairs of points is calculated. The pair with the smaller distance between the points has a greater weight.

The mentioned weight is given by the ratio of distances (appropriate significant points) from the selected point in the left image. Using the procedure described above, we can obtain better results than using simple averaging of possible positions and using the simpler rule proposed in [136]. The process of finding corresponding points by using simple averaging is briefly described in the following paragraph.

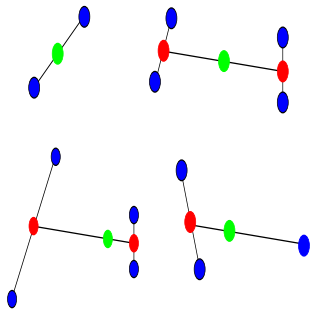


Fig. 2.12: Possible scatter of points and the process of calculating the final position of the point in the right image. Blue marks represent initial positions, red marks represent interim results and green marks represent the final position of the point in the right image. The final position is calculated as the progressive weighted average of initial positions. Weight is given by the distance between points in pairs.

The first two steps are the same as in the previous procedure. We find a set of corresponding points by SURF and we select five near points using the calculation of Euclid distance. Subsequently, we calculate differences between the depths of individual near points and their distances from the selected point. In the case that the ratio of depths exceeds a chosen threshold, the position of a selected pixel point in the right image is calculated only from the two closest points. Influence of each point is given by the ratio of its distance from the selected point. Otherwise, the

position of the point is obtained by averaging displacements of the nearest five points.

2.3.3 Experiments and results

The method described in section 2.3 was proposed to solve the problem with finding correspondences. We executed the tests for confirming its functionality. For executing the experiment, the proposed algorithm was implemented to the application (system) described in chapter 2.5. Finding the corresponding point was executing in the 6 six different images during the test. The obtained results were compared with results obtained using the method SAD (sum of absolute differences). The obtained results are assessed by the divergence from the positions determined accurately by the operator. Therefore, the reference position for comparison is given by manually determination. The summary of the achieved results are in Tab. 2.3. The table can be divided to three parts. The part of table named Accurate position contains the true position of the point in the right image. The position of the point in the left image is not stated because it does not have any meaning in evaluation of the test. The second part includes the Euclidean distance between true position and the position obtained by our proposed method and the method SAD is used for comparison. The part of table named Objective properties contains parameters which characterize the neighborhood of the corresponding point in the left image. We determine three parameters in the neighborhood 3x3: entropy, correlation and standard deviation. Subsequently, we evaluate the influence of the objective properties on the result. The value of the correlation obviously does not have a relationship with the accuracy of finding corresponding points. We can observe certain dependency of the results on the Entropy. The results are commonly more accurate if the entropy is higher. The strongest impact on the results has standard deviation. There is direct proportionality that accuracy of the results increased (euclidian distance decreases) with growing standard deviation. The influence is less significant for our proposed methods. The dependency is plotted in Fig. 2.14. The results of the experiment confirm our assumptions and feasibility of the proposed method. Obviously, the common method failed when the neighborhood of the selected points is dull and featureless. The proposed method allows to obtain better results in this situation. The average value of standard deviation for good results is 6.0762. On the contrary, the average value for wrong results is 4.712. Another interesting fact and significant information is that the average euclidian distance from the true positions is, for our proposed method 4.69, whereas for the method SAD it is 32.99. The advantage of using the proposed methods is obvious from the results. Moreover, this chapter contains a graphical expression of the results obtained by the proposed algorithm.

The results are represented in two ways. The first way is visualization in the form of a spatial model in Fig. 2.15. The second way is showing of the position of the corresponding points in the left and right image. The figure contains the position of the near significant points in both images and potential positions of the corresponding point in the right image besides the positions of the corresponding point. The input left and right images used for method verification are shown in Fig. 2.13.

Accurate position		Euclidian distance		Objective properties		
Vertical position	Horizontal position	Proposed method	SAD	Entropy	Correlation	Standard deviation
669	62	2.9101	3.6056	5.2950	138.87	0.2231
875	1187	0.6692	0	6.2494	172.40	0.0919
845	738	7.0318	12.2066	5.2745	111.60	0.04104
640	1150	0.9133	1	6.6407	224.80	0.1884
730	534	1.0688	27.5136	4.5316	18.47	0.0240
407	238	24.9141	53.4603	5.7739	25.81	0.0908
255	275	3.4759	3	6.8113	52.33	0.1967
406	233	0.9057	1	5.9948	47.10	0.0983
412	273	1.4936	30.8707	3.8570	101.53	0.0697
422	210	1.2492	1.4142	6.7961	51.11	0.1467
222	389	3.1427	1	5.3990	166.99	0.2190
297	104	1.1664	13	3.2600	313.75	0.0098
342	72	8.5518	150.2132	5.51	107.19	0.093
220	262	18.7277	31.9061	5.18	144.47	0.0384
909	354	0.9220	2.2361	6.4600	167.78	0.1773
1590	370	0.9054	1	5.9607	202.12	0.0918
1520	1518	3.6982	46.0109	4.4988	31.97	0.0249
1490	2096	7.7466	126.0357	4.2472	24.17	0.0190
1316	2164	1.6452	2	6.5052	231.00	0.135
605	1436	2.7529	152.3286	4.6887	174.14	0.0271

Tab. 2.3: Comparison of finding corresponding points by the proposed method and SAD and the influence of the properties of the point vicinity.

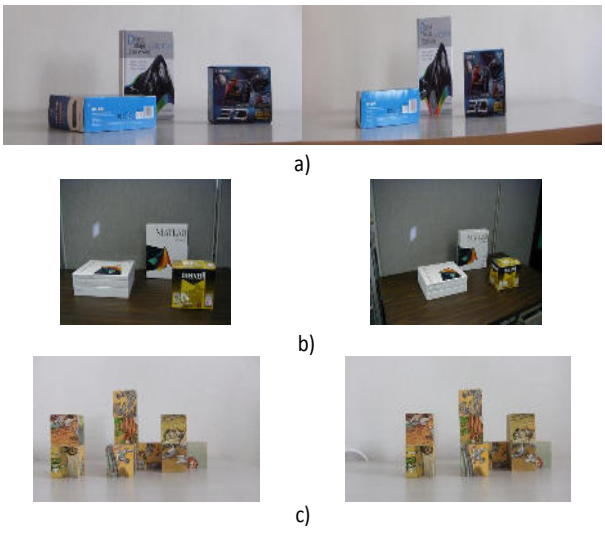


Fig. 2.13: Left and right input images used for method verification a) Boxes scene b) MATLAB scene c) Cubes scene.

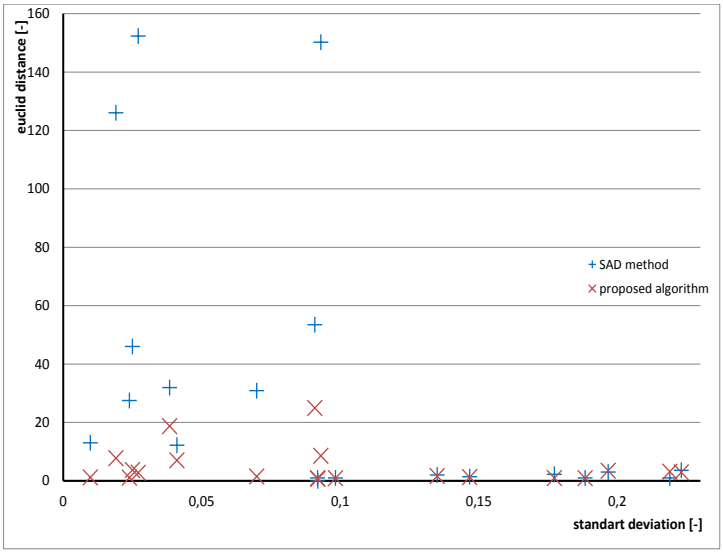


Fig. 2.14: Dependency of the accuracy (represented by euclidean distance from the accurate results) of finding corresponding points on the standard deviation.

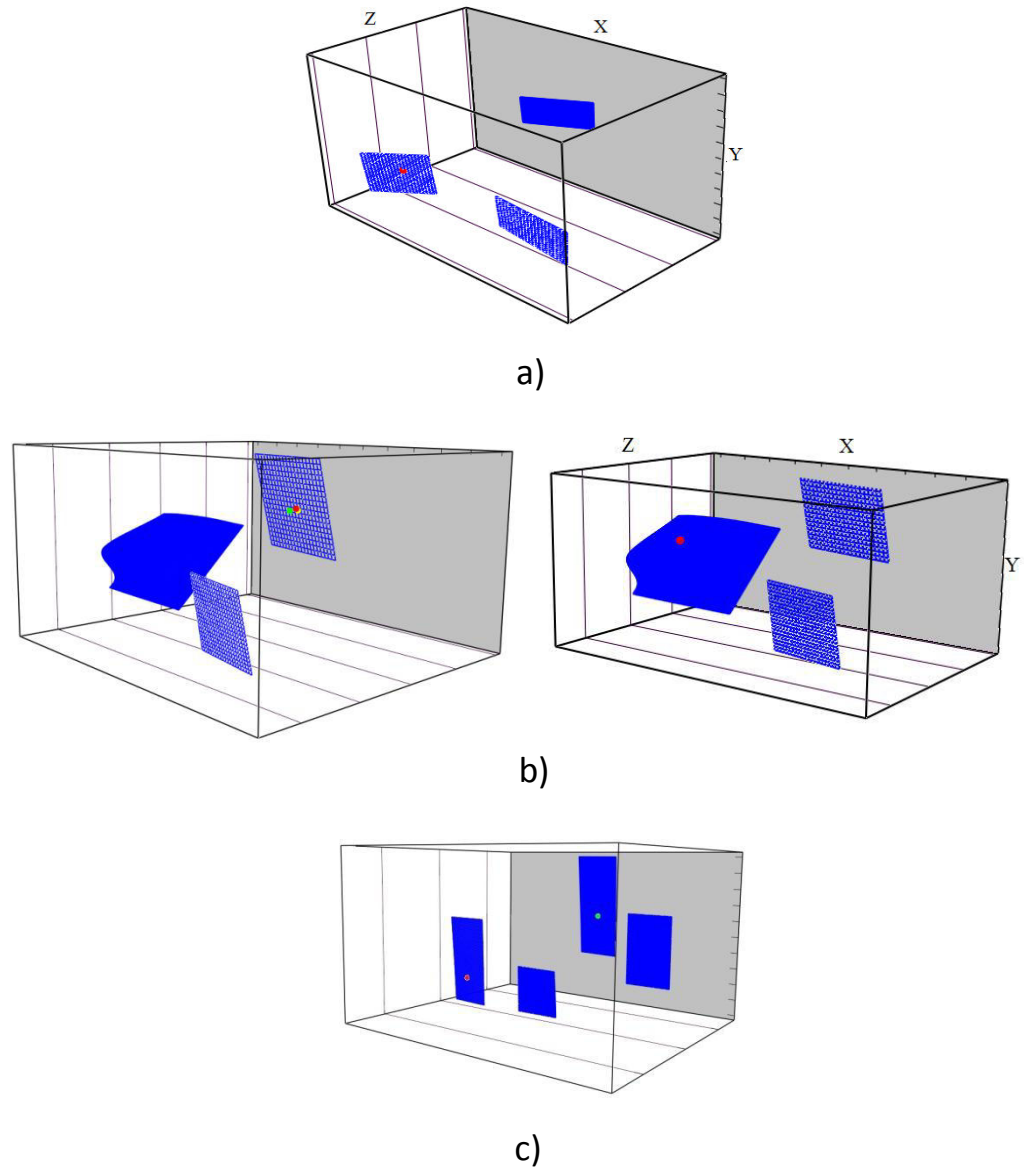


Fig. 2.15: Resulting position of reconstructed points. Red marks represent locations of selected points in space. Blue objects are pictured only for clarity. Model of a) Boxes scene b) MATLAB scene c) Cubes scene.

2.4 Utilizing the image in pseudo-color

2.4.1 Fundamental idea

The importance of a high-quality set of corresponding points for quality reconstruction is obvious from preview chapters. It is necessary to satisfy two requirements for creating this accurate set. Sufficient reliability of corresponding points is the first of them. Suitable spatial distribution of the corresponding point in the image is the second demand. Covering all objects in the scene is important for evaluating the sufficiency of the spatial distribution. The suitable distribution of corresponding points across the whole scene is a fundamental problem. The spatial distribution of corresponding points is partly given by number of finding corresponding points and partly by properties of the image. The number of corresponding points is influenceable. The properties of each object– brightness, contrast, standard deviation etc., are important aspects. Finding correspondences by the commonly used method based on similarity is problematic in areas with regular texture or with monotone brightness. This problem can be partly solved by the method proposed in chapter 2.3. Another possibility is making an effort to increase the contrast in problematic areas without correspondences. The use of pseudo- colors is one of the ways to perform this aim.

Pseudo- coloring is a technique for converting gray scale images to false colors (pseudo- colors), that do not correspond to real colors of the scene. This method allows a significant increase of contrast in the scene. The main application of the false (pseudo) color is for human analysis, because humans recognize more color levels than grayscale degrees. Pseudo- coloring is used mostly in biomedical applications (eg. [96],[97]). However, this method is applied even in other areas such as security (eg [98], [99]) or in mining missing data [100].A pseudo- colored image is described by three color components, as well as true RGB images (with true colors) or HSV. This section deals with the possibility of using pseudo- coloring for finding corresponding points in areas without contrast. This approach is novel. In previous research, we investigated the possibility of using pseudo- color space for image registration [132].

Another advantage is the possibility of increasing contrast especially in areas where we need it. The process of pseudo- coloring combines increasing space dimension and scale transformation. Brightness transformation can be used in areas with low contrast. The efficiency of using pseudo- color increases with using preliminary analysis of the input image. The parameters of the conversion can be adjusted based on this analysis. The disadvantage is the increase of computational complexity caused by multidimensional space.

2.4.2 Used methods

Different methods for converting an image to pseudo- coloring have been published. We can find a survey of some of them in [133]. We implemented some of these methods in our experiment. The first used method defines the conversion using parametric equation of curve in RGB space [101] (further called Color curve). This method published by Thomas M. Lehmann ensures a lot of color changes which can allow to find corresponding points. The method maintains the original progression of brightness. This method is based on the fact that pseudo- coloring can be described mathematically by a transformation curve in color space. The curve is equidistantly sampled to create as many points as there are input gray values. Each gray value is mapped to the specific color defined by the coordinates of the corresponding sample point in the color space.The method can be described by the following math equations.

$$\begin{pmatrix} R \\ G \\ B \end{pmatrix} = \frac{1}{2\sqrt{3}} \begin{pmatrix} 1+\sqrt{(3)} & 1-\sqrt{(3)} & 2 \\ 1-\sqrt{(3)} & 1+\sqrt{(3)} & 2 \\ -2 & -2 & 2 \end{pmatrix} \begin{pmatrix} r(t)sin(\omega t+\varphi) \\ r(t)cos(\omega t+\varphi) \\ z(t) \end{pmatrix}, \tag{2.57}$$

where t is the input value of the pixel in grayscale, φ the initial color (color of pixel with zero brightness) and ω specifies the dynamics of color changes.

These parameters are inputs to the conversion. Fig. 2.16 shows a model of the conversion with varying parameter ω. The marks represent positions of the resulting pixel values belonging to each gray scale level. The space represents a RGB cube. Through changing these parameters, we can affect the conversion; hence we can affect the output image. Consequently, we can influence the search of correspondence by changing these parameters. Moreover, results are influenced by the orthogonal distance $r(t)$ between the spiral curve and the main diagonal. The functions $r(t)$ and $z(t)$ must detain the curve in the finite range of the RGB-cube. We can derive that $r(t)$ and $z(t)$

$$r(t) = \sqrt{\frac{3}{2}} \begin{cases} t & if 0 \leq t \leq 0.5 \\ (1-t) & elsewhere. \end{cases} \tag{2.58}$$

A detailed description of a method with the relationship deriving of the relations is in [101].

The second method is based on color space. A color space is a mathematical representation of our visual perceptions. The frequently used color spaces are RGB and HSV (Hue , Saturation, Value). We used a method based on HSV. In HSV

space, a gray scale image f(x, y) can be represented as

$$\begin{gathered} V = f(x,y) \\ H = \frac{2\pi f(x,y)}{L} \\ S = \begin{cases} k \cdot f(x,y) & f(x,y) \leq \frac{L}{2} \\ k(L - f(x,y)) & f(x,y) < \frac{L}{2} \end{cases}, \end{gathered} \tag{2.59}$$

where L is the maximal gray levels of f(x, y), k is a constant factor (usually k = 1.5) and H, S, V are components of the color model HSV. Then, the pseudo-color transform can be performed by converting HSI into RGB color space using the following relations [103] [102]

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1 & -0,204124 & 0,612372 \\ 1 & -0,204124 & -0,612372 \\ 1 & 0,408248 & 0 \end{bmatrix} \begin{bmatrix} I \\ V_1 \\ V_2 \end{bmatrix}, \tag{2.60}$$

where

$$\begin{cases} V_1 = S \times cos(H) \\ V_2 = S \times sin(H) \end{cases}. \tag{2.61}$$

2.4.3 Implementation

In the previous section, the fundamental idea and used methods were described. Subsequently, we need to investigate applicability of the idea in various scenarios with various level of the reality. In the first level of reality, the same images are considered. Fig. 2.17 shows that with pseudo- color imaging, it is possible to find corresponding points in image areas where it was impossible in grayscale image (see Fig. 2.18). The search for corresponding points works perfectly in pseudo- colors when corresponding pixels have exactly the same brightness value in both gray scale images. Such a condition is ensured if both images (picture, photograph) have been captured at the same time with the same light conditions and with the same CCD sensor. Otherwise, a problem appears, because differences in pixel values increase due to the process of pseudo- coloring. Two approaches can be used for solving this problem.

The first approach is eliminating false (incorrect) correspondences using some restrictions which are deduced from reliable correspondences found in a monochromatic image. The proposed algorithm for elimination was created as an extension of the algorithm published in [95]. The extension is executed by adding more restrictions. The rules combine the restriction of the horizontal parallax, extreme in the angles of the line connecting corresponding points and similarity of the neighborhood of the examined image point. The restricted conditions are obtained from

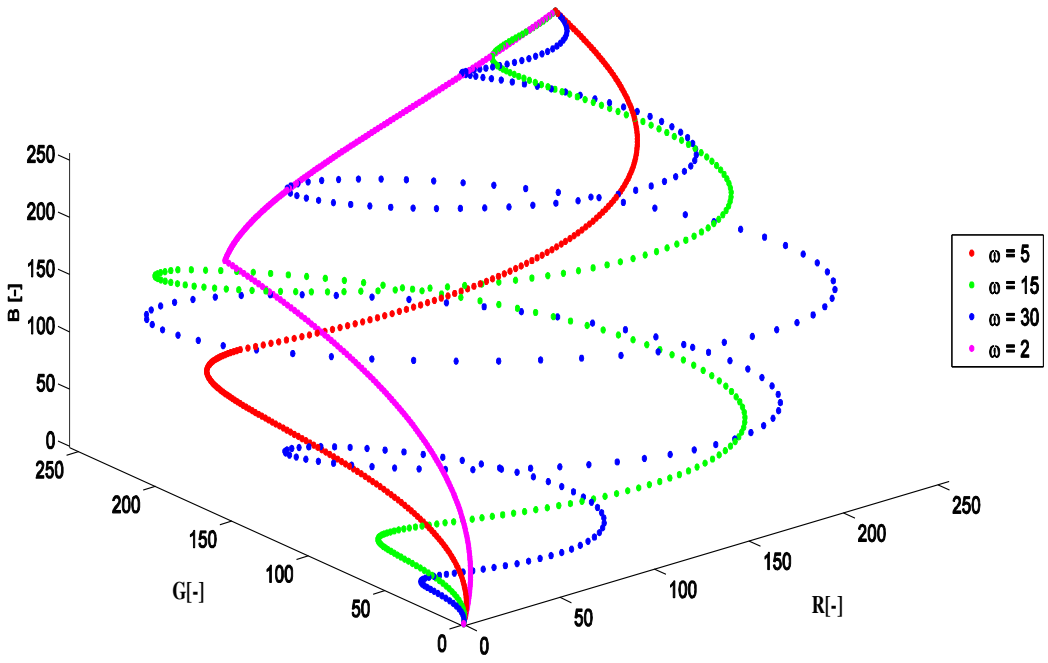


Fig. 2.16: The positions of the resulting pixel values belonging to each gray scale level. The space represents a RGB cube. The conversion was executed by Color Curve method with various parameters ω.

the set of reliable correspondences in the first step. Subsequently, the obtained conditions are compared with properties of the correspondences found in pseudo- color. We deal with angle conditions first. The average angle β is calculated. The β is an angle between the line connecting corresponding points in both images and horizontal axes. The allowed interval is calculated for the angle. The correspondence is evaluated as false in the case that an angle of the appropriate correspondences does not lie in this interval. The second restriction is the restriction of horizontal parallax. The parallax between the left and right image have to be positive in the real image. Subsequently, the maximal possible value of parallax is deduced from the values of the parallax of the reliable correspondences. The last restriction is the constraint of the difference of the brightness values in the defined neighborhood of the appropriate corresponding point. At first, the average difference between the neighborhoods of reliable correspondences is calculated. Subsequently, the difference between neighborhoods of the corresponding points found in pseudo- colors is compared with the obtained average value. However, the difference is calculated

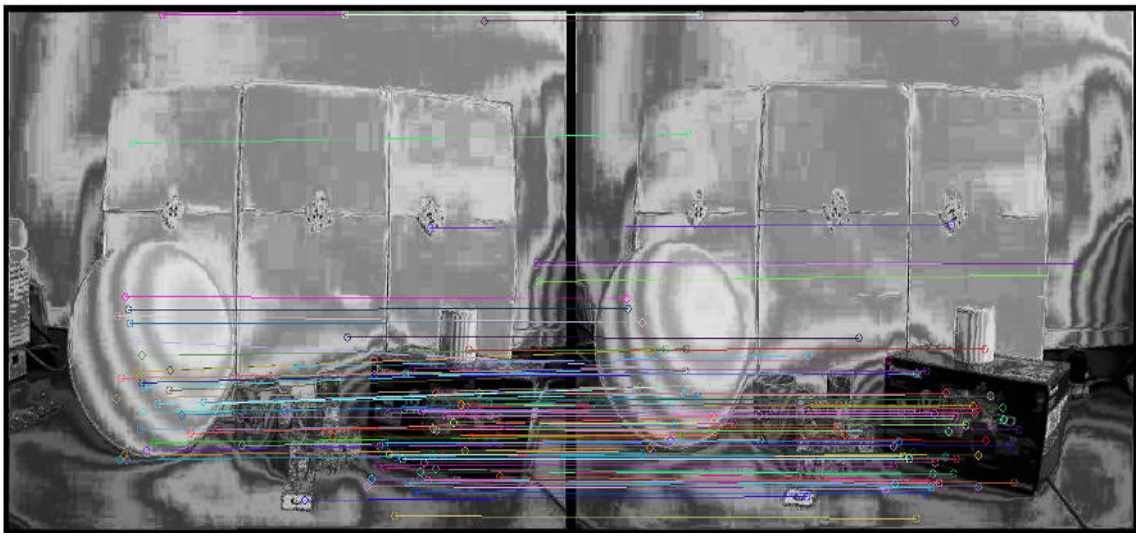


Fig. 2.17: The correspondences found in the pseudo-color image (shown in gray scale for better clarity).

from the values in the relevant positions of the monochromatic images. If the difference for the appropriate correspondence exceeds the predefined threshold, then the correspondence is identified as false.

The second approach uses methods for image enhancement in gray scale for better results. In this method, we transform the scale from one picture to another with the aim to eliminate the difference between the pixel values (brightness) in the corresponding pictures before converting them to pseudo- color. A suitable transformation can be deduced from the relation of the brightness value of the corresponding points found in the gray scale image.

Practically, in the proposed approach, pseudo- coloring is used only in the case where it is not possible to find enough corresponding points in gray scale images (by the method described above in Section 3.2). In such a case, we do not have sufficiently close points (found by SURF) for reliable results and we need to add some extra information. Then, we convert only the neighborhood of the selected point. Subsequently, we can find correspondences using SURF in the pseudo-color image. Finally, we use the algorithm to estimate the position of the selected point in the right image using the acquired correspondences. With respect to computational complexity, only converting the relevant part(in which we need to obtain correspondences) of the image is advantageous. Using pseudo- coloring has some risks without support of reliable correspondences found in the gray scale image.

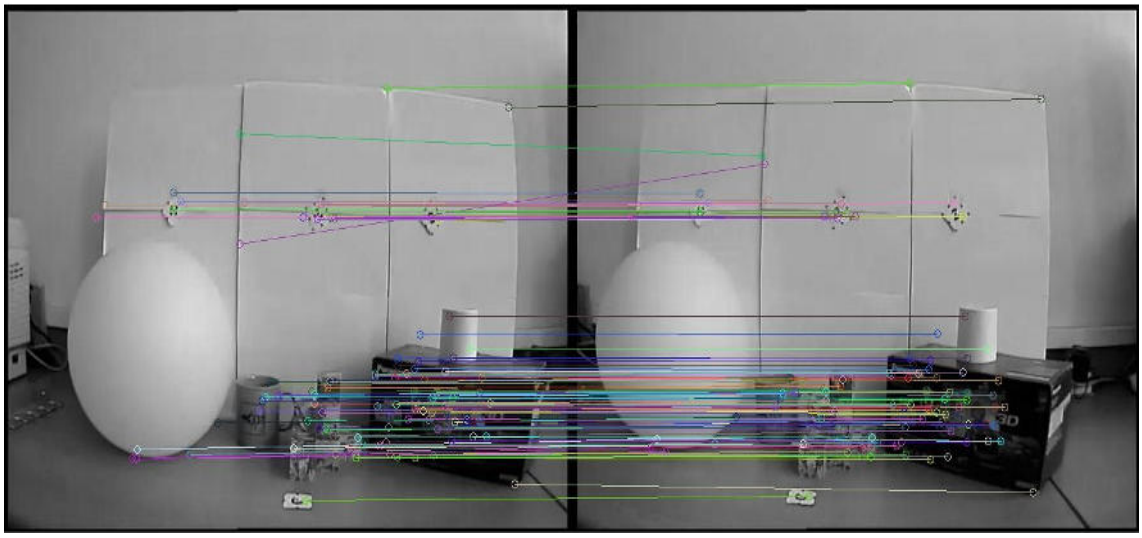


Fig. 2.18: The correspondences found in the monochromatic image.

2.4.4 Results

In this section, comparison of using a gray scale image and pseudo- colors image are presented. The greatest advantage of conversion to pseudo color is finding corresponding points in areas where it is impossible to find them in a gray scale image. This described advantage is demonstrated in the figures. Few correspondences (less than in pseudo colors with a higher threshold) in gray scale were found only after extremely decreasing the decided threshold (from 1000 to 50). This decreasing of the threshold increases the reliability of the correspondences too. The disadvantage of this method is sometimes the decreasing reliability. Increasing the reliability is possible by the approach described above. Moreover, the new approach was found during testing. The approach, further called SumPC, uses the sum of each component of the pseudo color and finding the correspondences in the obtained representation. In this test we used various methods for pseudo- coloring. The procedure of the tests was the following:

- input images were converted by different methods to pseudocolor,
- finding points corresponding in monochromatic image, in true color image and in various pseudo- color images,
- calculating disparities,
- comparing disparities with disparities of the particular pixel true depth map,
- calculating the number of errors and reliability (ratio of the error and number of found correspondences).

The test was executed with nineteen images from the Middbury Stereo Database [89]. The resolution of the images is 1110x1350. The results are summarized in Tab 2.4. Conversion to pseudo color by various methods and with various parameters was used in the experiment. The average reliability of the correct correspondences found in pseudo color is comparable with the reliability of finding correspondences in grayscale. The best results (reliability 91.37%) were reached when correspondences were searched in the grayscale image obtained by the sum of individual components of pseudocolored images.

Representation	**Method**	**Parameter**	**Reliability** [%]
Gray scale	origin	-	84.53
Pseudo Color	Color curve	$\omega = 5°$, $\phi = 5°$	79.33
Pseudo Color	Color curve	$\omega = 15°$, $\phi = 5°$	86.14
Pseudo Color	Color curve	$\omega = 30°$, $\phi = 5°$	84.61
Pseudo Color	based on HSV	k=1.4	80.48
Pseudo Color	based on HSV	k= 14	85.83
Gray scale	sumed based on HSV	k=4	91.37

Tab. 2.4: Average reliability of finding corresponding points in various representations of an image in a set of images from database (see Fig. 2.4) [89].

2.5 Designed software: Implementation of the proposed approach

In previous chapters were described the proposed methods for finding corresponding points in two images. In another chapter, we investigate achievable accuracy of the reconstruction and a method for estimating depth map. Designing an application was suitable. The application served for research. A graphical user friendly interface was designed, therefore the application can be used for practical purposes or as an educational tool. The advantage of the application is the possibility of selecting from more methods in the most offered procedure. In the application, some known approaches and open source solutions are used, besides the proposed methods. The application allows the following procedures:

- finding significant points, estimation of their correspondences between partial images,
- image rectification,
- interior calibration of the camera,

- exterior calibration of the camera,
- calculating spatial coordinates of the select point,
- reconstruction of a spatial model of the scene,
- generation depth map of a scene from two or more images.

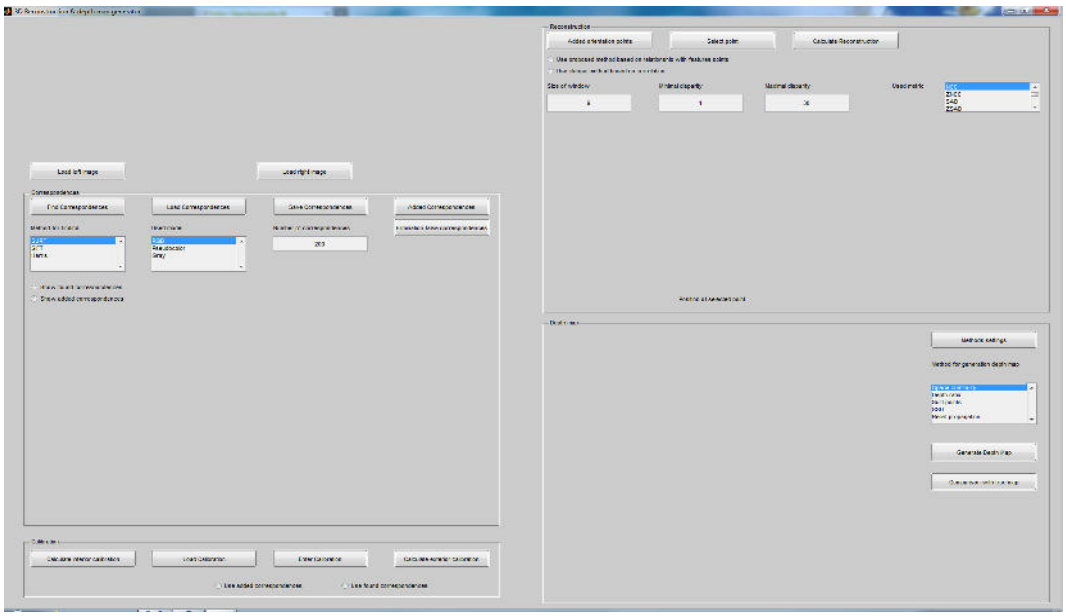


Fig. 2.19: The user interface of the created application.

The interface is shown in Fig. 2.19. The workplace is subdivided to five base sections. Going from upper left, there is a section for loading input images and displaying them. Below this, we can see a section dealing with correspondences. The last section on left side relates camera calibration. The section allowing reconstruction is positioned in the upper right. The section serving for estimating the depth map is in the bottom left.

2.5.1 Finding corresponding points

A significant part of this doctoral dissertation is devoted to finding corresponding points, subsequently the application deals with this task in large scale. There is a large offer of possibilities. The application offers the following few methods for finding corresponding points: SURF, SIFT, Harris detector and also the Fast Radial Feature detector. Two of these methods use point descriptor and the other two are only detectors. The descriptor submits some information about found points which serves to easier determine correspondences. The comparison of the performance of the implemented method is examined in section 2.2.4. the description of the method is in sections 2.2.1-2.2.3. The application allows to find corresponding points in

images in various interpretations. Besides monochromatic gray scale images and RGB images which are commonly used, the application used the image in pseudo colors and in the HSV color model. Using pseudo- colors is proposed and described in chapter 2.4.1. The next offered possibility is eliminating false correspondences by the proposed method. The application also allows to establish correspondences manually. The user, of course, can set the required number of correspondences. The elimination of false correspondences is possible using our proposed method or by the known algorithm RANSAC. The user can load their own correspondences saved on a disk. Furthermore, there is the possibility to save new found correspondences. Both loaded and saved correspondences are in MATLAB format *.mat.

2.5.2 Camera calibration

After establishing correspondences, camera calibration can be performed. The procedures used during camera calibration are described in chapter 2.1. The interior calibration matrix can be assigned in three ways. The first possibility is manually entering in the case that user knows it. The second way is loading file from a disk. The calculation of the calibration matrix is the last possible way. For this purpose, we used a function from open source Calib toolbox [83]. The toolbox is built based on the algorithm published by Heikkila [84]. Subsequently after obtaining calibration matrix $\mathbf{K}$, the user can perform exterior calibration. The input to this process is a set of corresponding points. The user is able to elect which set they want to use:

- points before eliminating false correspondences,
- points after eliminating false correspondences,
- manually added points,
- found points.

Subsequently, all data necessary for reconstruction is available.

2.5.3 The reconstruction of the spatial coordinates and spatial model

The main object is to obtain spatial coordinates of the individual points, the total model of the scene in the image and subsequently estimating the depth map. The application offers a few functions. The first possibility is calculating the overall model (button: Calculate reconstruction). In this case, spatial coordinates of the found corresponding points are calculated. The spatial model is shown and the user can manipulate it. The second possibility is calculating spatial coordinates for a specific selected image point. In this case, the user selects (by clicking) on

a point in the left image and the function will find its corresponding point in the right image. The resulting spatial positions of the selected points are graphically expressed in the model and in the front view. Moreover, their numeric value is written. The corresponding point can be found by one of the classic locally based methods: NCC, ZNCC, SAD, ZSAD, LSAD, SSD, ZSSD, LSSD, GRAD or their combination. If classic methods are elected, then the user sets parameters: size of the window, and minimal and maximal disparity. The corresponding point can also be found by our proposed method based on the relationship with feature points. The application allows to add extra orientation points which can help in subsequent finding correspondences.

2.5.4 Estimating the depth map

The last section of the application window allows to estimate the depth map. Even in this part of the created application, the user can use various methods to obtain the depth map. The interface contains check boxes and edit boxes which the user uses for accurate specification for creating the depth map. Various procedures are implemented in the application which can be mutually combined. The used algorithms are described in section 4.1. The edit boxes allow to set the value of parameters used in procedures. The user can choose default settings of the procedures. This way is very helpfull for inexperienced users. On the other hand, extensive setting options are beneficial for users with experience or for education, and also if we want to obtain various depth maps for research. Of course, the application served for testing the impact of these parameters.

3 ACCURACY OF THE METRIC RECONSTRUCTION ANALYSIS

The accuracy of reconstruction is a very important issue in the practical use of the reconstruction. Various aspects affecting accuracy of the reconstruction are discussed in this chapter. Establishing and comprehending the significance of each source of error are very important. We focused mostly on the impact of the error incurred during the finding corresponding points on the other steps in the process of determining spatial coordinates of a particular point and consequently on the error in the resulting determination of the spatial coordinates. This fact is in accordance with the topic of this dissertation work (see sections 2.2, 2.3). The theoretical part was supported by executed experiments. We have to distinguish between two various situations. In normal situations, the calculation of space coordinates is executed using the following equations [81]

$$Z = f_c \cdot \left(\frac{B}{x_2 - x_1} \right), \tag{3.1}$$

$$X = \frac{Z \cdot x}{f_c} = \frac{B \cdot x}{x_2 - x_1}, \tag{3.2}$$

$$Y = \frac{Z \cdot y}{f_c} = \frac{B \cdot y}{x_2 - x_1}, \tag{3.3}$$

where X, Y and Z are spatial coordinates of the points, x_{imi} and y_{im} are image coordinates of the points, B is stereo base and f_c is focal length of both cameras.

In the general case, the procedure from section 2.1 is used for the calculation. The errors of the spatial coordinates are different in these two situations due to various mechanisms of the calculation. Therefore, these situations will be solved separately. This section is divided into two sub- sections depending on the following aspects. The influence of incorrect determination of corresponding points' positions is investigated in section 3.1. The section 3.2 deals with influence of incorrect camera alignment (exterior calibration). The camera system parameters B and f_c also influence the accuracy of the reconstruction. The analysis of this influence belongs to a comprehensive analysis of the accuracy. However, this issue is described in detail in literature [46]. Therefore, in this dissertation, this analysis is not executed.

3.1 The influence of correspondence error points

Some error always occurs in spatial reconstruction, even if accurate calibration is assumed. The error is caused by small errors in determining corresponding points.

The first test ensures the basic for other advanced experiment. Firstly, the stereo case is briefly described. The equations for error in stereo alignment of the cameras can be derived from equations 3.1, 3.2 and 3.3. The equations for error in spatial coordinates assuming correctly determined focal length and stereo base are the following [81]

$$\Delta Z = \frac{Z}{f_c}\frac{Z}{B}\delta_{px}, \tag{3.4}$$

$$\Delta X = \sqrt{\left(\frac{x_1}{f_c}\frac{Z}{f_c}\frac{Z}{B}\delta_{px}\right)^2 + \left(\frac{Z}{f_c}\delta_x\right)^2}, \tag{3.5}$$

$$\Delta Y = \sqrt{\left(\frac{y}{f_c}\frac{Z}{f_c}\frac{Z}{B}\delta_{px}\right)^2 + \left(\frac{Z}{f_c}\delta_y\right)^2}, \tag{3.6}$$

where $\delta_{px}, \delta_x, \delta_y$ represent error of determining the parallax and image coordinates, therefore, it expresses errors in correspondences. ΔX, ΔY and ΔZ represent errors in spatial coordinates. The resulting error also depends on the two ratios $\frac{Z}{f_c}$ and $\frac{Z}{B}$, besides error in correspondences. Consequently, the errors increase with increasing depth Z of the point.

The error in the general case is a more complex problem. Therefore, the estimation of the error is obtained directly by calculating of the error in a particular situation. The calibration matrix $\mathbf{K}$, rotation matrix $\mathbf{R}$, translation vector $\mathbf{T}$ and correct spatial positions of the 2675 corresponding image pairs are known. The left and right images of the real scene used in experiment scanned by various camera systems is shown with found corresponding points in Fig. 3.1 [104]. The model of the scene is in Fig. 3.3, where blue marks represent correct positions of the points and red markers represent reconstructed positions. The spatial positions of the points (data on the axes) are related to the position of the first camera, therefore the focal of the camera is the coordinate center.

The dependencies of the error in all spatial coordinates and overall error of the point's position on the Z position are shown in Figs. 3.5- 3.8. The term overall error is used. The overall error is defined as a percentage expression of the ratio of two distances. The first of them is the euclidean distance of the correct and incorrect reconstructed position and the second one is the distance between the center of the coordinate system and the correct reconstructed position. The absolute errors are shown in Fig 3.1. The relative values are obtained from the ratio with the correct point's coordinates. The analysis of the accuracy of the reconstruction was executed for three various camera alignments. The errors for various camera alignments are plotted by various colors. The errors increase with increasing depth. This fact is consistent with the situation in the stereo case parallel optical axes. However,

the curves of the dependencies are not constantly increasing (specially the error in coordinate Z) because the error depends on more conditions. The depth increases with decreasing horizontal parallax in stereo alignment and this fact is one of reasons for increasing depth error. However, in the general case, this dependency may not be valid. Figure 3.4 shows the relation between horizontal parallax and depth for three different camera alignments.

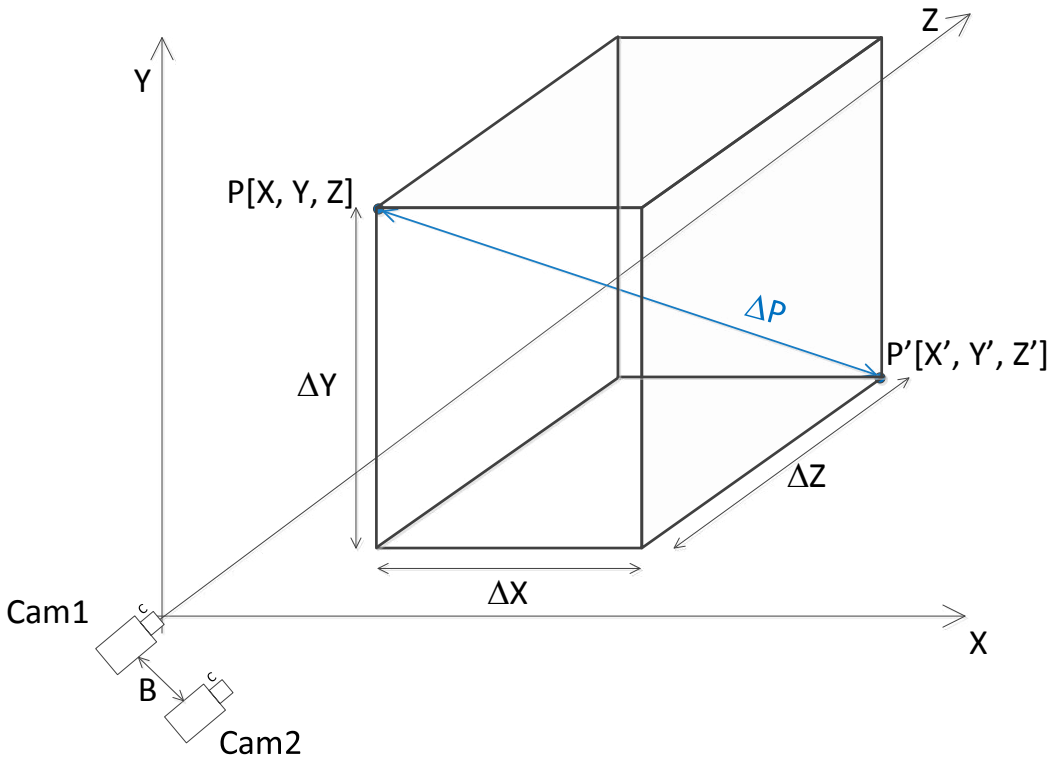


Fig. 3.1: The illustration of the absolute error in spatial coordinates including overall error ΔP. The coordinate center is located in the optical center of the first camera.

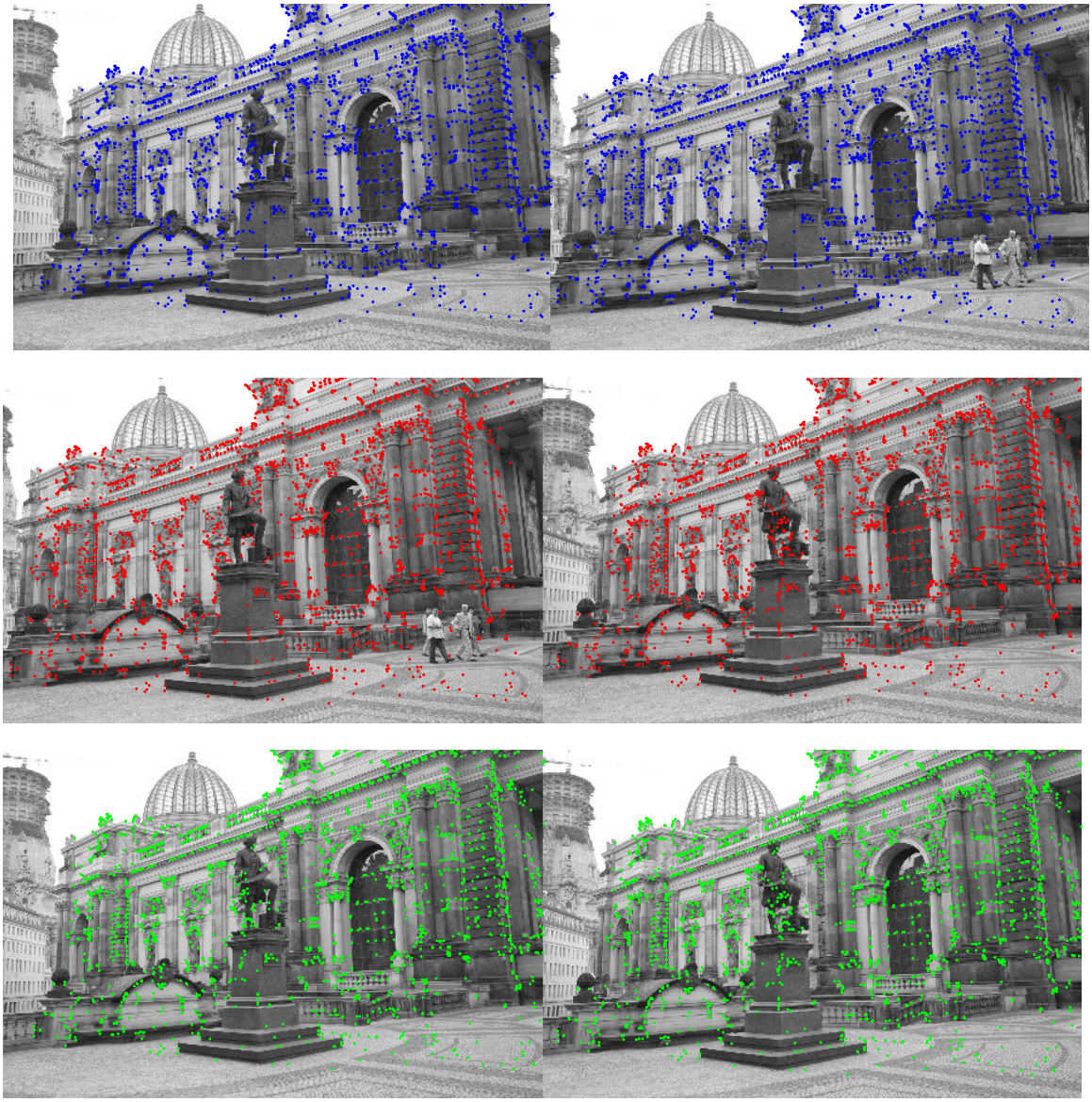


Fig. 3.2: The three pairs of images of the same scene [104] captured by various cameras systems. Significant points used in the basic test are marked in the scene. Points reconstructed by different systems are marked by various colors. The same color is used in the following Figs. 3.4- 3.8 to distinguish errors for various camera systems.

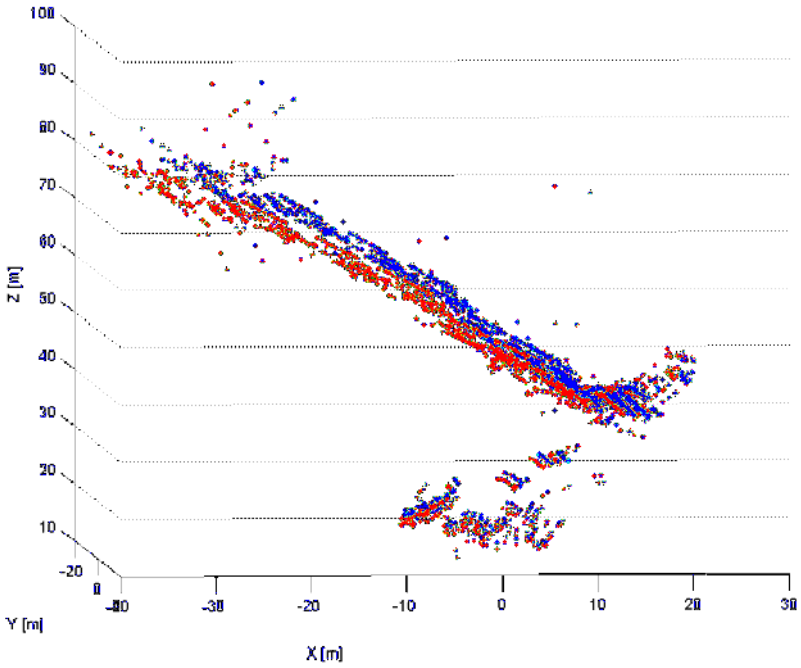


Fig. 3.3: The model of the scene, blue marks represent represent positions of the points and red markers represent reconstructed positions.

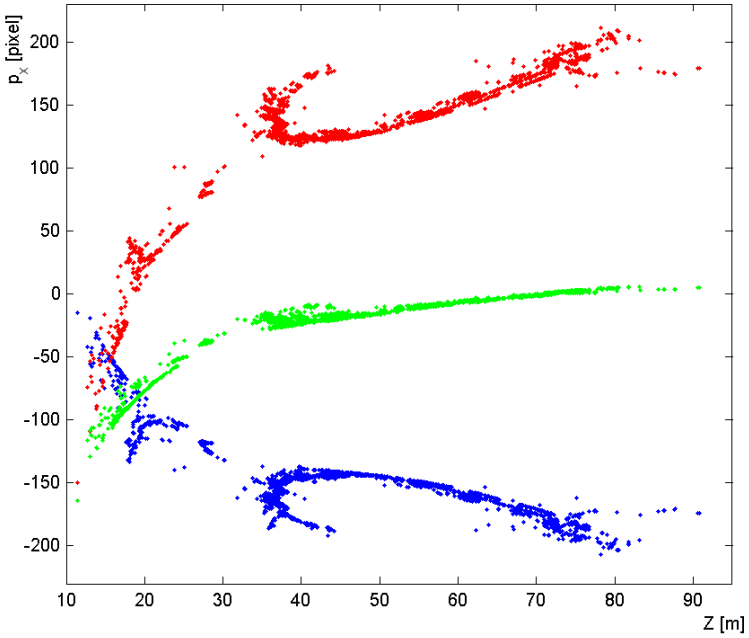


Fig. 3.4: The dependency of the horizontal parallax p_x on the depth Z of the point for three different camera systems captured scene 3.1. Points reconstructed by different camera systems are marked by various colors in conformity with the color marking in Fig. 3.1.

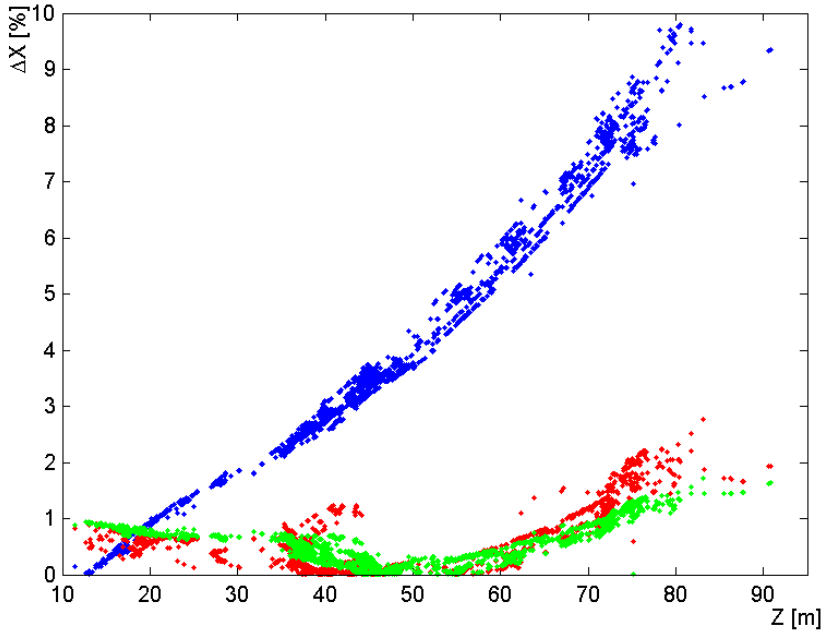


Fig. 3.5: The dependency of the relative error ΔX of the horizontal space coordinate X on the depth coordinate Z for three different camera systems captured scene 3.1. Points reconstructed by different camera systems are marked by various colors in conformity with the color marking in Fig. 3.1.

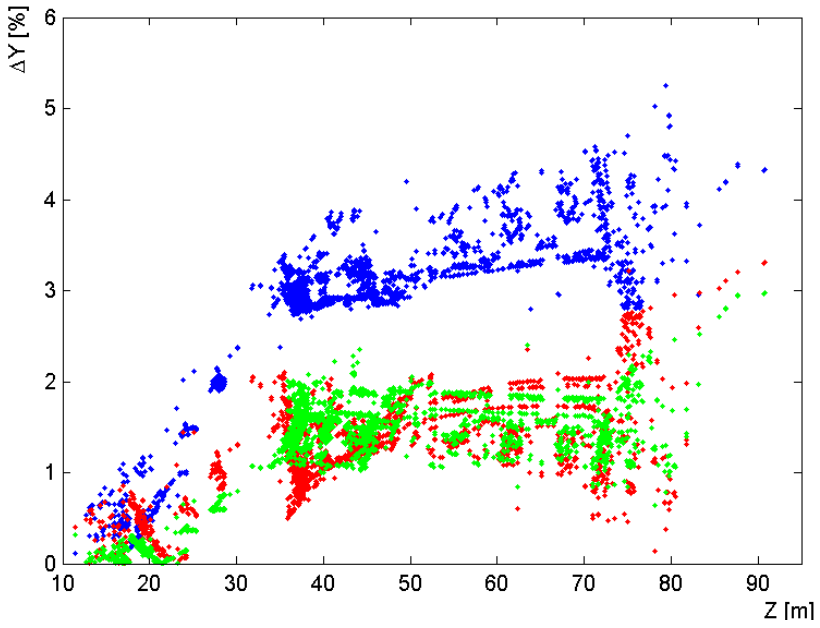


Fig. 3.6: The dependency of the relative error ΔY of the vertical space coordinate Y on the depth coordinate Z for three different camera systems captured scene 3.1. Points reconstructed by different camera systems are marked by various colors in conformity with the color marking in Fig. 3.1.

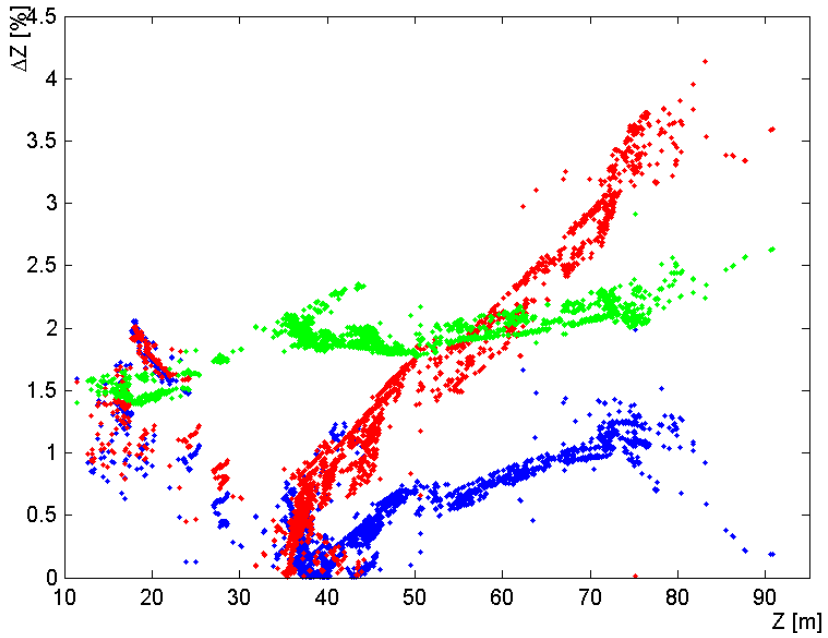


Fig. 3.7: The dependency of the relative error ΔZ of the depth space coordinate Z on the depth coordinate Z for three different camera systems captured scene 3.1. Points reconstructed by different camera systems are marked by various colors in conformity with the color marking in (see Fig. 3.1).

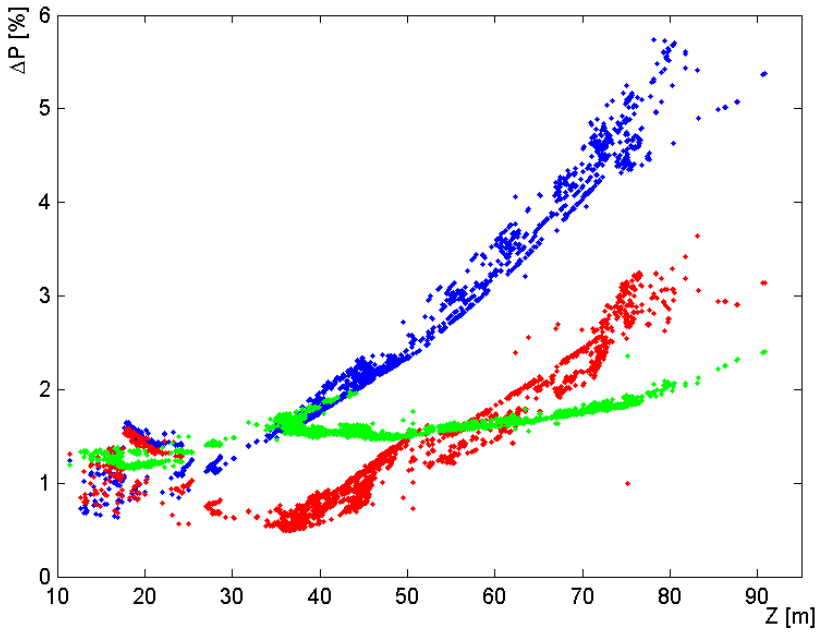


Fig. 3.8: The dependency of the overall relative error ΔP of the space position on the depth coordinate Z for three different camera systems captured scene 3.1. Points reconstructed by different camera systems are marked by various colors in conformity with the color marking in (see Fig. 3.1).

3.2 The influence of inaccurate camera alignment

The accuracy of the reconstruction depends on the geometry of the cameras, especially on the correctness of its determination. The normal case (stereoscopic) can be considered as the basic state of the camera system. The camera system can be transformed from the general case to the basic state by using the projection matrix. Then, various errors in camera alignment can occur in the camera system in the basic (stereo) state. The errors in camera rotation are investigated in this work, these error are represented by the error angles α, β and γ. The situation is shown in Fig. 3.9. There are two various practical situations which can be represented by these errors. In the first case, the cameras were originally in normal positions with parallel optical axes. Then, angles represent real error in physical position of the cameras. In this case, if camera alignment is correct, positions of the corresponding points differ only in horizontal position, if cameras alignment is right. However, this assumption is not valid if cameras are not in perfect normal position. Therefore, the corresponding points cannot be found if we suppose a accurate stereoscopic system, because corresponding points are being searched only in the same row.

In the second case, the cameras were originally in the general positions and were transformed to the stereoscopic state by using the projective matrix obtained during exterior calibration. The calibration can be wrong and then the angles represent error in the calibration. The geometry is represented by the projection matrix and determined by interior and exterior calibration of the cameras (see 2.1.3). Therefore, the camera calibration is a very important step in determining the spatial coordinate from stereo images in the viewpoint of accuracy. In this section, we will examine the error in determining the space coordinates depending on incorrectly establishing mutual camera position. Imperfect exterior calibration affects finding corresponding points and vice versa. We determine exterior calibration of the camera using found corresponding points. Therefore, incorrectly determined corresponding points cause wrong determination of exterior calibration. Subsequently, wrong exterior calibration causes incorrect reconstruction of space coordinates.

We analyzed the effect of various errors in camera alignment on the system accuracy. This topic is based on article [49] and dissertation [51]. The authors derived a formula for errors ΔZ from geometric situations. However, we executed practical experiments and discovered that the derived formula is simplified and is valid only in special situations when the point lies on the horizontal axes of the image. Therefore, we executed a new analysis of the situations. Consequently, the results achieved by both derived formulas were compared. Subsequently, I extended the original analysis published in this paper by examining the error in all three space coordinates. Moreover, the relation between these error angles and error in

finding corresponding points was investigated. The set of the found corresponding points serves for obtaining the projective matrix, which represents information about camera alignment. We assume a basic stereoscopic system described above in section 2.1 or in [81].

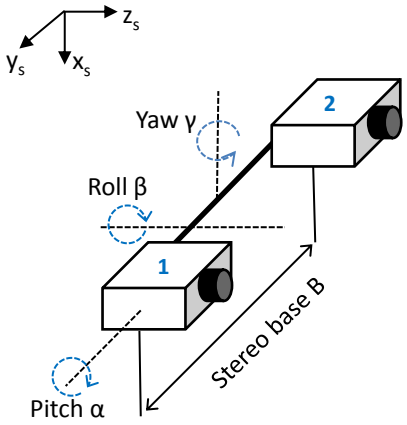


Fig. 3.9: Normal scanning system with two cameras with marking of possible fault angles α, β, γ.

3.2.1 Errors in stereo positions of the cameras

The coordinate system is modified due to the error in camera alignment. Consequently, image coordinates of the spatial points are changed. At first, we derive a general formula for calculating errors in all spatial coordinates in dependency on the incorrectly determined image coordinates. The relation for error in depth ΔZ is derived in [49]. Formulas for errors in the other two dimension ΔX and ΔY are derived in this chapter. In the next step, the relations for incorrectly determined image coordinates x_2' and y' are found. The relations are derived from geometrical situations by using trigonometric functions. Therefore, these relations are various for various calibration errors. Subsequently, this relation is substituted into the general relation.

The following equations were obtained by using formulas describing simple stereophotogrammetry (3.1)[49]

$$Z_{error} = f_c \cdot \left(\frac{B}{x_2 - x_1} - 1 \right) \simeq f_c \cdot \left(\frac{B}{x_2 - x_1} \right), \tag{3.7}$$

$$Z_{real} = f_c \cdot \left(\frac{B}{x_2' - x_1} - 1 \right) \simeq f_c \cdot \left(\frac{B}{x_2' - x_1} \right), \tag{3.8}$$

where Z_{error} is the observed absolute depth from image plane to the object. Z_{real} is the real(true) absolute depth, f_c is the focal length of both cameras (we assume

the simple case, where the cameras are the same), B is stereo base (length of the base line), x_1 is the correct (true) position of the measured point in the first image obtained by the first camera, x_2 is the true position of the measured point in the second image obtained by the second camera and x'_2 is the error (observed) position of the particular pixel in the second image captured by the real second camera.

The vertical image coordinate y_{im} can also be changed. This change brings a problem which [49] does not consider. The vertical image coordinate y_{im} is not included in the equation. However, a change of the vertical position can cause the corresponding point not to be found. An analysis of this problem in particular situations will executed. The error is calculated as the difference of the real and observed depth of the point.

$$\Delta Z = Z_{real} - Z_{error}. \tag{3.9}$$

Then,the following mathematical operations are executed

$$\begin{aligned} \Delta Z &= f \cdot \left(\frac{B}{x'_2 - x_1} \right) - f \cdot \left(\frac{B}{x_2 - x_1} \right), \\ \Delta Z &= f_c \cdot B \frac{x'_2 - x_2}{(x'_2 - x_1)\,(x_2 - x_1)}, \\ \Delta Z &= Z_{real} \left(\frac{x'_2 - x_2}{x'_2 - x_1} \right). \end{aligned} \tag{3.10}$$

The formula (3.10) represents the general error of the spatial coordinate Z. The coordinate x'_2 is substituted in the next step. The formula for coordinate x'_2 is derived for each situation (error in three various angles).

Subsequently, we deal with the derivation of the general formula for ΔX. The procedure is very similar to the derivation formula for ΔZ. The following equations are obtained by using the stereophotogrammetric formula (3.2)

$$X_{error} = B \cdot \left(\frac{x_2}{x_2 - x_1} - 1 \right) \simeq B \cdot \left(\frac{x_2}{x_2 - x_1} \right), \tag{3.11}$$

$$X_{real} = B \cdot \left(\frac{x'_2}{x'_2 - x_1} - 1 \right) \simeq B \cdot \left(\frac{x'_2}{x'_2 - x_1} \right), \tag{3.12}$$

where X_{error} is the observed absolute spatial horizontal coordinate. X_{real} is the correct real (true) absolute spatial horizontal coordinate. Other terms represent the same parameters as in equation (3.7).

Then ΔX can be expressed as

$$\Delta X = \frac{B(x_2(x'_2 - x_1) - x'_2(x_2 - x_1))}{(x_2 - x_1)(x'_2 - x_1)}. \tag{3.13}$$

Subsequently, we deal with the derivation of the general formula for ΔY. The procedure is very similar to the derivation formula for ΔZ. The following equations are obtained by using formula 3.3

$$Y_{error} = B \cdot \left(\frac{y}{x_2 - x_1} - 1 \right) \simeq B \cdot \left(\frac{y}{x_2 - x_1} \right), \tag{3.14}$$

$$Y_{real} = B \cdot \left(\frac{y'}{x'_2 - x_1} - 1 \right) \simeq B \cdot \left(\frac{y'}{x'_2 - x_1} \right), \tag{3.15}$$

where Y_{error} is the observed absolute spatial vertical coordinate. Y_{real} is the real (true) absolute spatial horizontal coordinate. Other terms represent the same parameters as in equation 3.7.

Consequently, ΔY can be expressed as

$$\Delta Y = \frac{B(y(x'_2 - x_1) - y'(x_2 - x_1))}{(x_2 - x_1)(x'_2 - x_1)}. \tag{3.16}$$

At this moment, all general equations necessarily required for expressing the errors have been expressed. Firstly we will deal with error in roll with rotation angle α between two cameras. We assume that first camera is perfect calibrated and its optical axis represents axis z of the coordinate system with center in the focus. Optical axes of the second camera are parallel to the optical axes of the first camera. However, second camera has wrong calibration. The error is in the angle α about optical axes. The geometric situation is shown in Fig. 3.10. Based on stereogrammetry, the following substitutions can be used [49]

$$x_1 = f_c \frac{X}{Z}, \tag{3.17}$$

$$x_2 = f_c \frac{X - B}{Z}, \tag{3.18}$$

$$y = f_c \frac{Y}{Z}, \tag{3.19}$$

where X_i and Y_i are the correct (true) 3D coordinate of the object and $X_2 - X_1 = B$. The author in [49] derived this substitution $x'_{2,Z} = x_2 \cdot cos(\alpha)$. Therefore, the following formula was obtained

$$\Delta Z = Z_{real} \frac{Y_2 \, (cos\alpha - 1)}{B}. \tag{3.20}$$

However, the error in the expression of $x'_{2,Z}$ can be proved. This formula is valid only if point P lies on the x axis. From Fig.3.10, it is obvious that x'_2 is equivalent to line segment $\overline{RS}$ while expression $x'_2 = x_2 \cdot cos(\alpha)$ used in [49] is equivalent to

line segment $\overline{RY}$. Obviously, $x'_2 = \overline{RY} - \overline{SY}$ and from triangle OSY $\overline{SY} = y \sin \alpha$ and therefore

$$x'_{2,P} = X \cos \alpha - Y \sin \alpha. \tag{3.21}$$

Similarly, y' can be derived from triangle OSY, where $y = \bar{OY}$ is the hypotenuse and from triangle RPY, where $x_2 = \bar{PY}$ is the hypotenuse. Therefore

$$y'_{2,P} = Y \cos \alpha + X \sin \alpha. \tag{3.22}$$

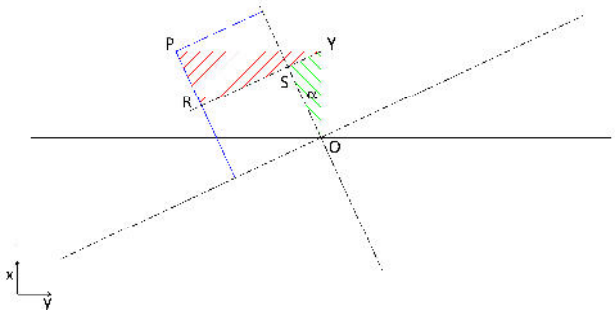
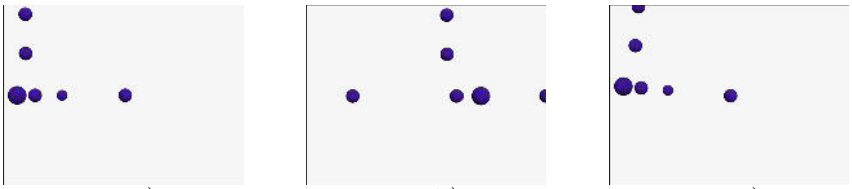


Fig. 3.10: The geometric situation for roll error.

The correctness of our derived formulas was proved experimentally. The whole following experiment was used for verifying of the correctness and will also be used for verifying the next error: pitch and yaw. Special software (3D CAD) for creating and rendering simple virtual scene was used in the experiment. The scene contains six spheres (see Fig. 3.2.1). The scene was rendered by cameras with accurately set parameters:

- translation between cameras so called stereo base B,
- rotation of the camera α, β, γ,
- focal distance f,
- sensor size d_S.

The examples of the rendered image are in Fig. 3.2.1. The experiment is based on finding particular points in the scene rendered by the left camera of the optimal stereo camera system x_2 and y_2. Then we compute the theoretical position of the point in the image obtained by rotated the left camera by using Zhao's formula $x'_{2,Z}, y'_{2,Z}$ and by our proposed formula $x'_{2,P}, y'_{2,P}$. Subsequently, the position of the point in the rotated image was found $x'_{2,F}, y'_{2,F}$. In the next step, the computed and found coordinates coordinates are compared. The results are in Tab. 3.1. It is obvious from the table that more accurate results are reached by using our newly derived formulas 3.21 and 3.22. The formulas proposed in [49] are only valid, if the vertical position is 0 (points lie on the vertical center of the image.)



a) b) c)

Fig. 3.11: Rendered image used for verifying of the formula for error in image coordinates a) left image without roll b) right image without roll c) left image with roll of the camera by 5°.

Consequently, formulas 3.21 and 3.22, which represent error in image coordinates caused by camera rotation, is successively substituted into formulas 3.10, 3.13 and 3.16, which represent general errors of spatial coordinates caused by error of image coordinates. The final equations for errors in all spatial coordinates are obtained and after math modification have forms

$$\Delta Z_T = Z_{real}\frac{X_2\,(cos\alpha - 1) - Y\sin\alpha}{B}, \tag{3.23}$$

$$\Delta X_T = \frac{BX}{X + B\cos\alpha - X\cos\alpha + Y\sin\alpha} - X, \tag{3.24}$$

$$\Delta Y_T = -\frac{XY + B^2\sin\alpha + Y^2\sin\alpha - BX\sin\alpha - XY\cos\alpha}{X + B\cos\alpha - X\cos\alpha + Y\sin\alpha}. \tag{3.25}$$

Tab. 3.1 also contains spatial coordinates of the given point computed by using the found position of the corresponding points

- positions of the points in the left and right camera in an ideal stereoscopic system $(X_{REAL}, Y_{REAL}, Z_{REAL})$,
- positions of the points in the left and right camera in a stereoscopic system with investigated error $(X_{ERROR}, Y_{ERROR}, Z_{ERROR})$.

The differences $(\Delta X_C, \Delta Y_C \Delta Z_C)$ between these spatial coordinates are computed by using formulas (3.26)-(3.28). Simultaneously, the theoretical error caused by rotation $(\Delta X_T, \Delta Y_T, \Delta Z_T)$ are computed by using formulas (3.23), (3.24), (3.25). Subsequently, theoretical and practical errors are compared. Obviously, these errors are equal. Therefore, the obtained relations can be used for estimating the error caused by the roll of the camera.

$$\Delta X_C = |X_{ERROR} - X_{REAL}|, \tag{3.26}$$

$$\Delta Y_C = |Y_{ERROR} - Y_{REAL}|, \tag{3.27}$$

$$\Delta Z_C = |Z_{ERROR} - Z_{REAL}|. \tag{3.28}$$

	pixel 1	pixel 2	pixel 3	pixel 4
x_1 [pixel]	56.00	134.00	79.00	46.00
y_1 [pixel]	113.00	0.00	0.00	111.00
x_2 [pixel]	263.00	298.00	227.00	128.00
y_2 [pixel]	113.00	0.00	0.00	111.00
$x'_{2,F}$ [pixel]	252.00	297.00	228.00	114.00
$x'_{2,Z}$ [pixel]	261.99	296.87	226.13	127.51
$x'_{2,P}$ [pixel]	252.15	296.87	227.66	112.85
$y'_{2,F}$ [pixel]	135.00	26.00	286.00	238.00
$y'_{2,Z}$ [pixel]	—	—	—	—
$y'_{2,P}$ [pixel]	135.45	25.97	284.92	237.92
X_{REAL} [mm]	-248.90	93.06	94.05	43.40
X_{ERROR} [mm]	-247.09	93.27	94.42	41.44
ΔX_C [mm]	1.80	0.22	0.37	0.12
ΔX_T [mm]	1.81	0.24	0.26	0.18
Y_{REAL} [mm]	-106.94	0.00	0.00	-104.70
Y_{ERROR} [mm]	-132.76	-18.10	16.97	-55.85
ΔY_C [mm]	25.82	18.10	17.97	-48.85
ΔY_T [mm]	25.83	18.06	17.99	-48.81
Z_{REAL} [mm]	1886.79	1385.68	2371.54	1886.79
Z_{ERROR} [mm]	1953.13	1389.32	2377.65	1801.80
ΔZ_C [mm]	66.33	3.64	6.11	84.99
ΔZ_T [mm]	64.58	3.64	6.19	89.83

Tab. 3.1: The verification of the proposed formulas (3.21), (3.22) for calculation error image positions $x'_{2,P}$, $y'_{2,P}$ and formulas (3.24),(3.25),(3.23) for calculation of the error of the spatial coordinates ΔX_T, ΔY_T, ΔZ_T for the roll of the camera.

Figure 3.13 illustrates relative error in depth (coordinate Z) in dependency on space coordinate X of the object. The error angle of the roll is a parameter of the curves. The error is related to depth (space coordinate Z). The error calculated by the formula proposed in [49] is plotted with dashed lines. The error calculated by the newly proposed formulas are plotted with solid lines. Figure 3.14 illustrates the relative error in the vertical space coordinate Y in dependency on horizontal space coordinate X of the object. The error angle of the roll is a parameter of the curves. The error is related to depth (space coordinate Z). Figure 3.12 illustrates relative error in the horizontal space coordinate X in dependency on horizontal space coordinate X of the object. The error angle of the roll is a parameter of the curves.

The error is related to depth (space coordinate Z). The error in coordinates X and Y was not considered in [49], therefore the comparison is impossible.

The dependencies on the stereo base B and space coordinates Y were investigated, however, it was not plotted on the graph. The errors increase with increasing Y and decreasing B. The increase of error with decreasing B complies with basic error in stereophotogrammetry; this dependency is typical for most phenomena in sterephotogrammetry. The absolute error increases with increasing depth.

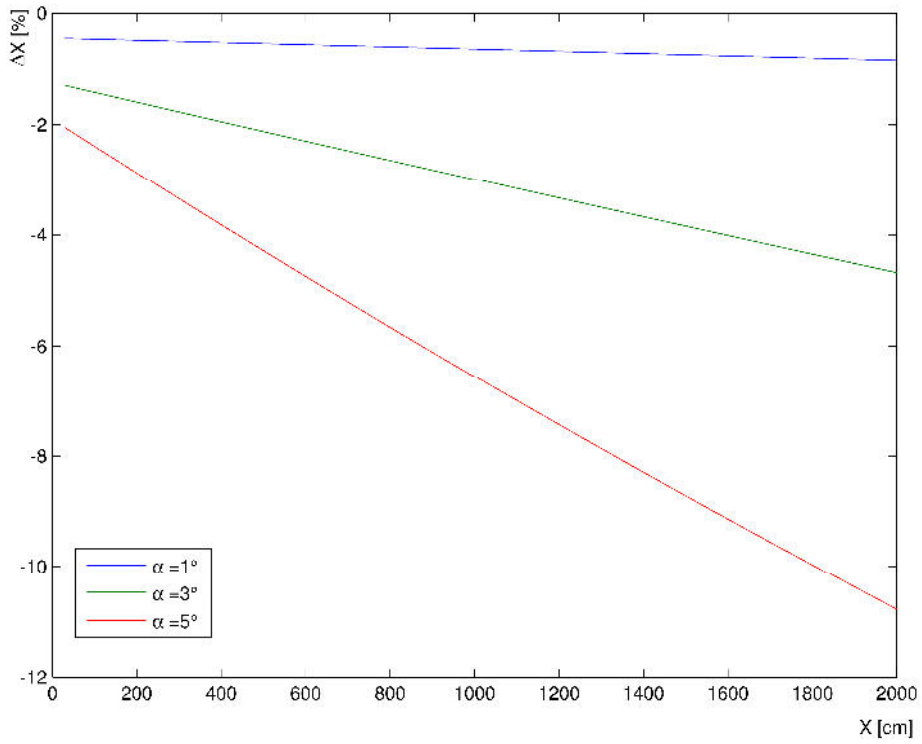


Fig. 3.12: The dependency of the relative error ΔX of the coordinate X on the roll angle α and space coordinates X. Used sensing system parameters B=75mm, f=8.5mm.

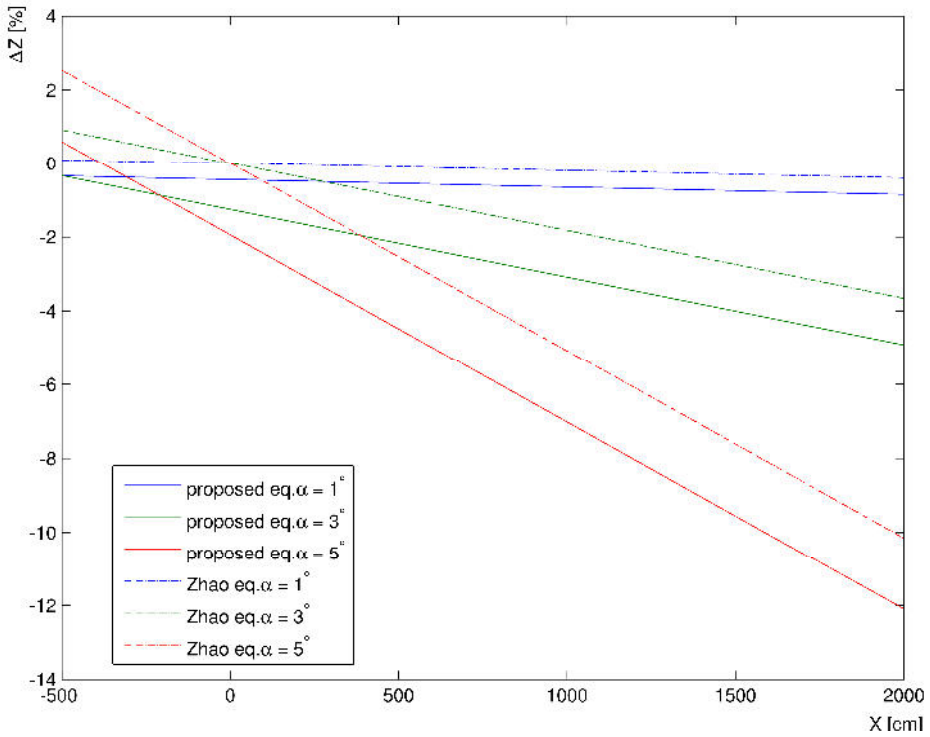


Fig. 3.13: The dependency of the relative error ΔZ of the coordinate Z on the roll angle α and space coordinates X. Used sensing system parameters B=75mm, f=8.5mm.

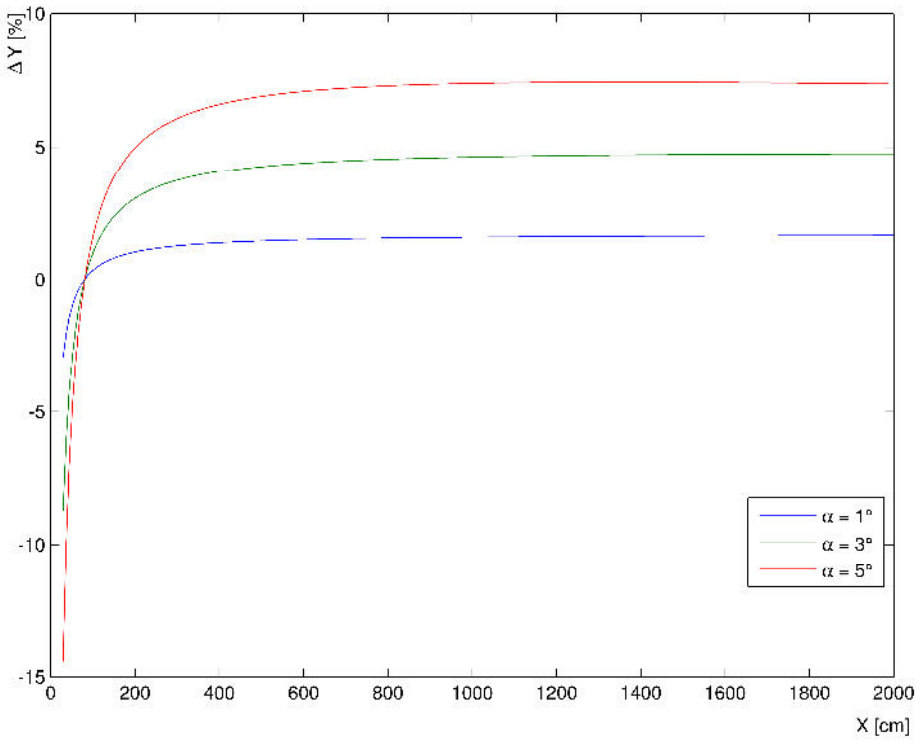


Fig. 3.14: The dependency of the relative error ΔY of the coordinate Y on the roll angle α and space coordinates X. Used sensing system parameters B=75mm, f=8.5mm.

Subsequently, we assume that the first camera is perfectly calibrated and aligned with the bar. The calibration of the second camera is perfect except for a certain rotation angle β about a line which is parallel to the bar. The stereoscopic geometry of this situation is illustrated in Fig. 3.16. An important fact is that the epipolar line is no longer parallel with the bar. Using trigonometry (see Fig. 3.16), the formula $y'_2 = y_2 \cdot cos(\beta)$ was derived in [49], where β represents pitch angle between two cameras. Subsequently, the author uses the fact that sec β is equal to $\sqrt{1 + (tan\beta)^2}$. Therefore $x'_{2,Z}$ can be obtained from x_2 and β by the following formula

$$x'_{2,Z} = \frac{x_2}{\sqrt{1 + (tan(\beta))^2}} \simeq x_2 \left[1 - \frac{1}{2}(tan\beta)^2\right]. \tag{3.29}$$

Angle β is usually considered as very small, therefore $tan\beta \simeq \beta$. Consequently, Taylor's series can be used to derive the right hand side of the above equation. Subsequently, y' is substituted to the general equation 3.10 and the formula for error in depth is obtained

$$\begin{aligned} \Delta Z &\simeq Z_{real} \frac{\left(x_2 \left[1 - \frac{1}{2}(tan\beta)^2\right] - x_2\right)}{y_2 - y_1}, \\ \Delta Z &\simeq \frac{1}{2} \frac{x_2 \left(Z_{real} tan(\beta)\right)^2}{fB}, \\ \Delta Z &\simeq -\frac{1}{2} \frac{X_2 Z_{real} tan(\beta^2)}{B}. \end{aligned} \tag{3.30}$$

Two various situations can be considered. Both cases are illustrated in Fig. 3.15.

- Situation I (Fig. 3.15 a): In this case $y'_2 = \frac{y_2}{cos(\beta)}$. Then the final equation for calculating the error is

$$\Delta Z \simeq Z^2_{real} \frac{x_2 \left(cos(\beta)^{-1} - 1\right)}{fB} \simeq Z_{real} \frac{X_2 \left(cos(\beta)^{-1} - 1\right)}{B}. \tag{3.31}$$

- Situation II (Fig. 3.15 b): The measurement error is 0.

The experiment described above in section 3.2.1 for verifying its correctness was executed. The experiment is again based on the found points position in rendered images. The experiment revealed that formulas (3.31) and (3.29) are in accordance with the real state. However, from the planar model used in [49], it is impossible to derive formulas for error in vertical image coordinate y. The spatial model of the situation was used to derive a more accurate formula (see Fig. 3.16). The basics of the derivation is finding the point of intersection of the plane ϑ and line segment $\bar{u}$ (denoted in Fig. 3.16). Firstly, the line segment $\bar{u}$ passes points $P[X, Y, Z]$ and $F'[f_c \cos\beta, 0, f_c \sin\beta]$. Then, the line segment $\bar{u}$ is described by the parametric equation

$$U = P - tF', \tag{3.32}$$

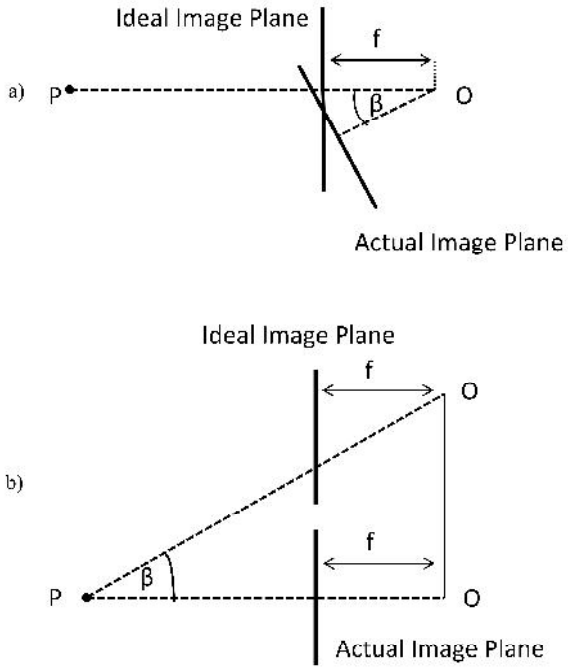


Fig. 3.15: Two special case of the error due pitch: (a) Type I (b) Type II.

The plane ϑ is described by using a general equation using three points, which lies on it $P_1[0,0,0]$, $P_2[0,1,0]$ and $P_3[\sin\beta, 0, \cos\beta]$.

$$0 = ax + by + cz. \tag{3.33}$$

Subsequently, the line segment equation is substituted to the plane general equation. Then parameter t is computed and substituted back to (3.45). After this substitution, the final position of x'_2 and y'_2 is obtained. After mathematical operations, the simplified formulas are obtained

$$x'_{2,P} = \frac{X f_c}{Z\cos\beta - f_c + Y\sin\beta}, \tag{3.34}$$

$$y'_{2,P} = -\frac{Y f_c \cos\beta^2 - Z f_c \cos\beta \sin\beta}{Z\cos\beta - f_c + Y\sin\beta}. \tag{3.35}$$

Subsequently, the experiment comparing Zhao's and the proposed formulas for calculating the change of image point position is executed. The results are in Tab. 3.2. It is obvious from the table that our newly derived formulas, 3.34 and 3.35, are usable. Consequently, these formulas are successively substituted to formulas 3.10, 3.13 and 3.16. The final equations for errors in all spatial coordinates are obtained

and after mathematical modification they have the following forms

$$\Delta X_T = X - \frac{BXZ}{XZ(1-\cos\beta) - Bf_c + Xf_c + BZ\cos\beta + BT\sin\beta - XY\sin\beta}, \tag{3.36}$$

$$\Delta Y_T = Y + \frac{B\left(Yf_c\left(\frac{\cos 2\beta}{2} + 0.5\right) - Zf_c\cos\beta\sin\beta\right)}{\left(\frac{Xf_c}{Z\cos\beta - f_c + Y\sin\beta + \frac{f_c(B-X)}{Z}}\right)(Z\cos\beta - f_c + Y\sin\beta)}, \tag{3.37}$$

$$\Delta Z_T = \frac{XZ^2}{B\left(Z\cos\beta - f_c + Y\sin\beta\right)} - \frac{XZ}{B}. \tag{3.38}$$

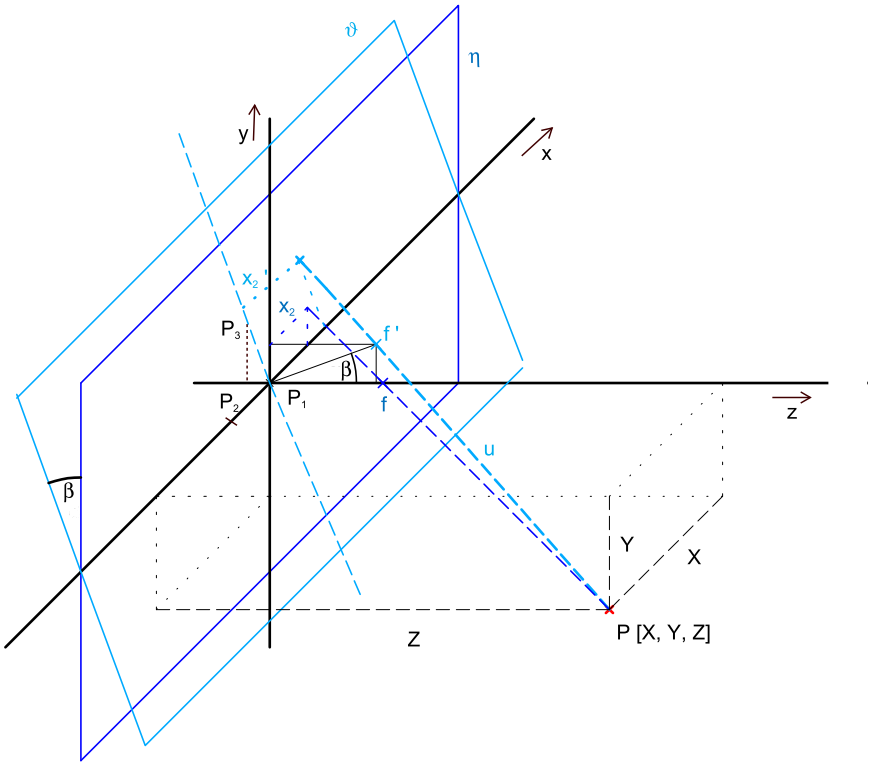


Fig. 3.16: The model of the geometric situation for pitch angle β. The dark blue plane represents the plane of the image without error. The skyblue plane represents the plane of the image with error. The formulas error of the image coordinates (3.34) and (3.35) are derived from this image.

Subsequently, the obtained formulas were verified using the same procedure as formulas (3.23), (3.24), (3.25). Therefore, the differences (ΔX_C, ΔY_C, ΔZ_C) between spatial coordinates in an ideal camera stereoscopic system (X_{real}, Y_{real}, Z_{real})

and a system with error in alignment (X_{error}, Y_{error}, Z_{error}) were calculated and compared with the theoretical error (ΔX_T, ΔY_T, ΔZ_T) obtained by the newly derived formulas (3.36), (3.37) and (3.38). The results for a few points are in Tab. 3.2. Theoretical errors and real differences are equal, and it is confirmed that the derived formulas are valid.

	pixel 1	pixel 2	pixel 3	pixel 4
x_1 [pixel]	607.00	471.00	452.00	612.00
y_1 [pixel]	300.00	163.00	197.00	160.00
x_2 [pixel]	399.00	263.00	298.00	404.00
y_2 [pixel]	300.00	163.00	197.00	160.00
$x'_{2,F}$ [pixel]	399.00	264.00	298.00	404.00
$x'_{2,Z}$[pixel]	399.93	263.08	298.50	403.99
$x'_{2,P}$[pixel]	399.01	264.00	299.04	403.97
$y'_{2,F}$[pixel]	368.44	233.00	268.20	230.00
$y'_{2,Z}$[pixel]	—	—	—	—
$y'_{2,P}$[pixel]	135.45	233.39	267.00	230.42
X_{REAL}[mm]	298.56	197.60	198.70	-5.77
X_{ERROR}[mm]	298,57	197.13	198.02	-5.73
ΔX_C[mm]	-0.01	0.47	0.68	0.04
ΔX_T[mm]	-0.01	0.47	0.70	0.04
Y_{REAL}[mm]	0.00	197.60	-783.12	201.92
Y_{ERROR}[mm]	-100.61	-96.52	644.04	-100.35
ΔY_C[mm]	100.61	294.11	1427.16	302.27
ΔY_T[mm]	100.61	294.11	1427.09	302.25
Z_{REAL} [mm]	2884.62	2884.62	3894,10	2884.62
Z_{ERROR} [mm]	2867.94	2897.85	3922.48	-2884.23
ΔZ_C [mm]	16.68	13.23	26.37	0.39
ΔZ_T [mm]	16.68	13.23	26.34	0.39

Tab. 3.2: The verification of the proposed formulas (3.34), (3.35) for calculation error image positions $x'_{2,P}$, $y'_{2,P}$ and formulas (3.36),(3.37),(3.38) for calculation of the error of the spatial coordinates ΔX_T, ΔY_T, ΔZ_T for the pitch of the camera.

Figures 3.17, 3.18, 3.19, illustrate dependencies of the relative error of the space coordinates X, Y, Z on the parallax p_x, horizontal x_{im} and vertical y_{im} image positions and on stereo base B. There are many various dependencies which can be investigated and plotted. It is possible to monitor dependencies on the nine input parameters: focal length f_c, stereo base B, horizontal image position x_{im} or

alternative horizontal space position X, vertical image position y_{im} or alternative vertical space position Y, parallax p_x or alternative depth space coordinate Z and error angle. Then it is possible to investigate: new error vertical image position, new error horizontal position and errors in three space coordinates. The plotting of all dependencies would be space-consuming. The relative errors are plotted because it more aptly informs us about error severity. Due to this fact, errors for small vertical position are not included in the graph because the relative error reaches a very high value as consequence of dividing by small the number. The coordinate Y is the most sensitive to error in pitch and its relative error reaches a value of about 20 5% for angle 1°. While, the error in the next two spatial coordinates reached values up to 5% for a given angle. Therefore, the error in image coordinate y_{im} has crucial importance for the accuracy and feasibility of calculating spatial coordinates.

The formula (3.16) for calculating the error in coordinate Y contains the vertical image coordinate y_{im} too. This error is not considered in article [49]. However, the error in vertical coordinate y_{im} is more significant for pitch than the error in horizontal coordinate x_{im}. This fact is obvious from the comparison in Tab. 3.2. Moreover, the most critical problem is feasibility of the calculation. Calculating the error assumes correctly finding corresponding points in the image obtained by the rotated camera. The corresponding points are found in the row in which lies in the first image. This means that the corresponding point is not found if the vertical image coordinate is changed due to rotation. This hypothesis is valid for all error in alignment of the camera system. Consequently, the corresponding points cannot be found by a simple algorithm working in one row if there is the assumption that error in alignment occurs. Therefore, pitch error is the most critical error from the view of findability of correspondences.

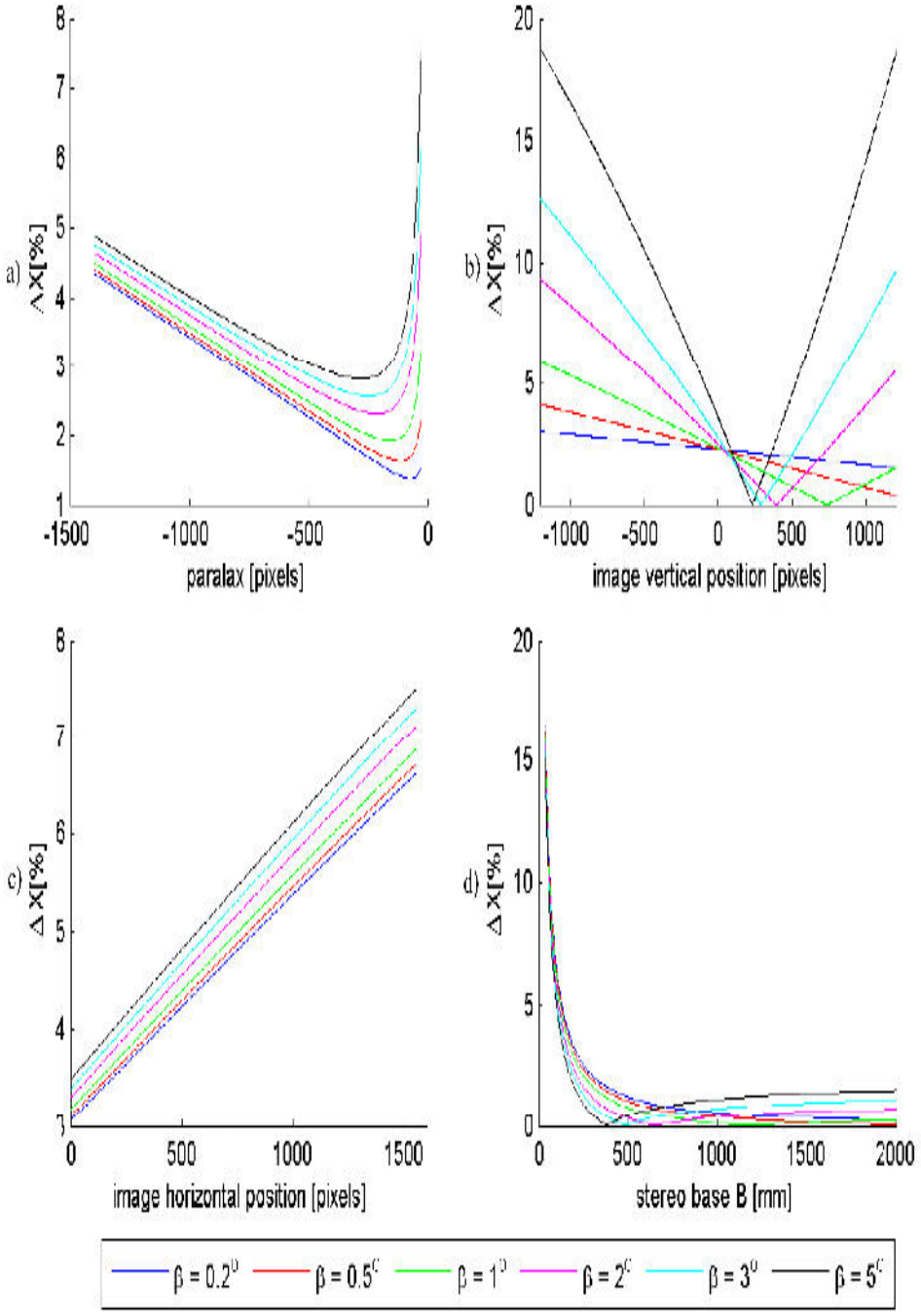


Fig. 3.17: The dependency of the relative error ΔX in the horizontal space coordinate X on the a) horizontal parallax b) image vertical position, c) image horizontal position, d) stereo base. The fault angle β is a parameter. Used parameters of the camera system B=500mm, f=8.5mm.

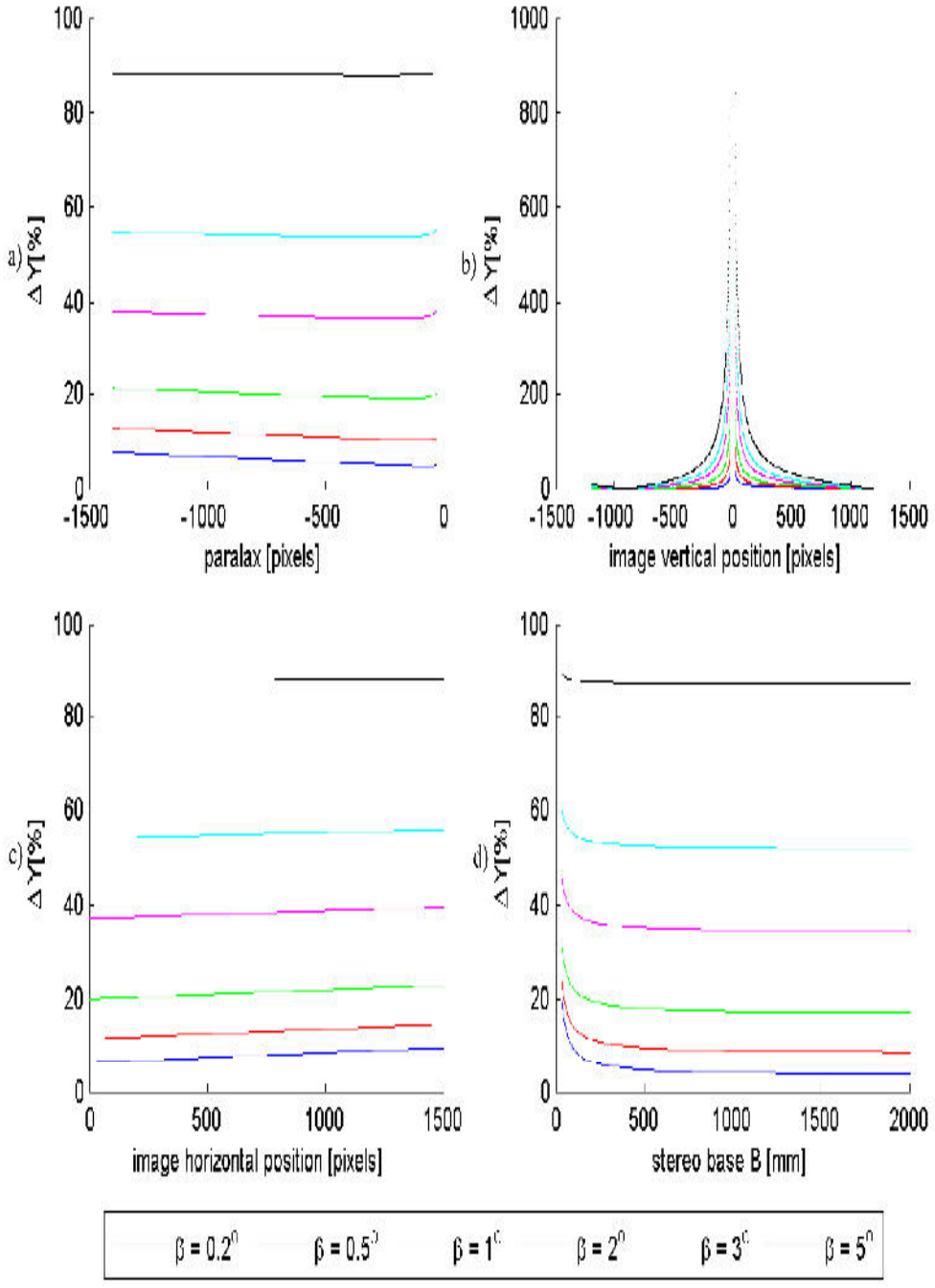


Fig. 3.18: The dependency of the relative error ΔY in the horizontal space coordinate Y on the a) horizontal parallax b) image vertical position, c) image horizontal position, d) stereo base. The fault angle β is parameter. Used parameters of the camera system B=500mm, f=8.5mm.

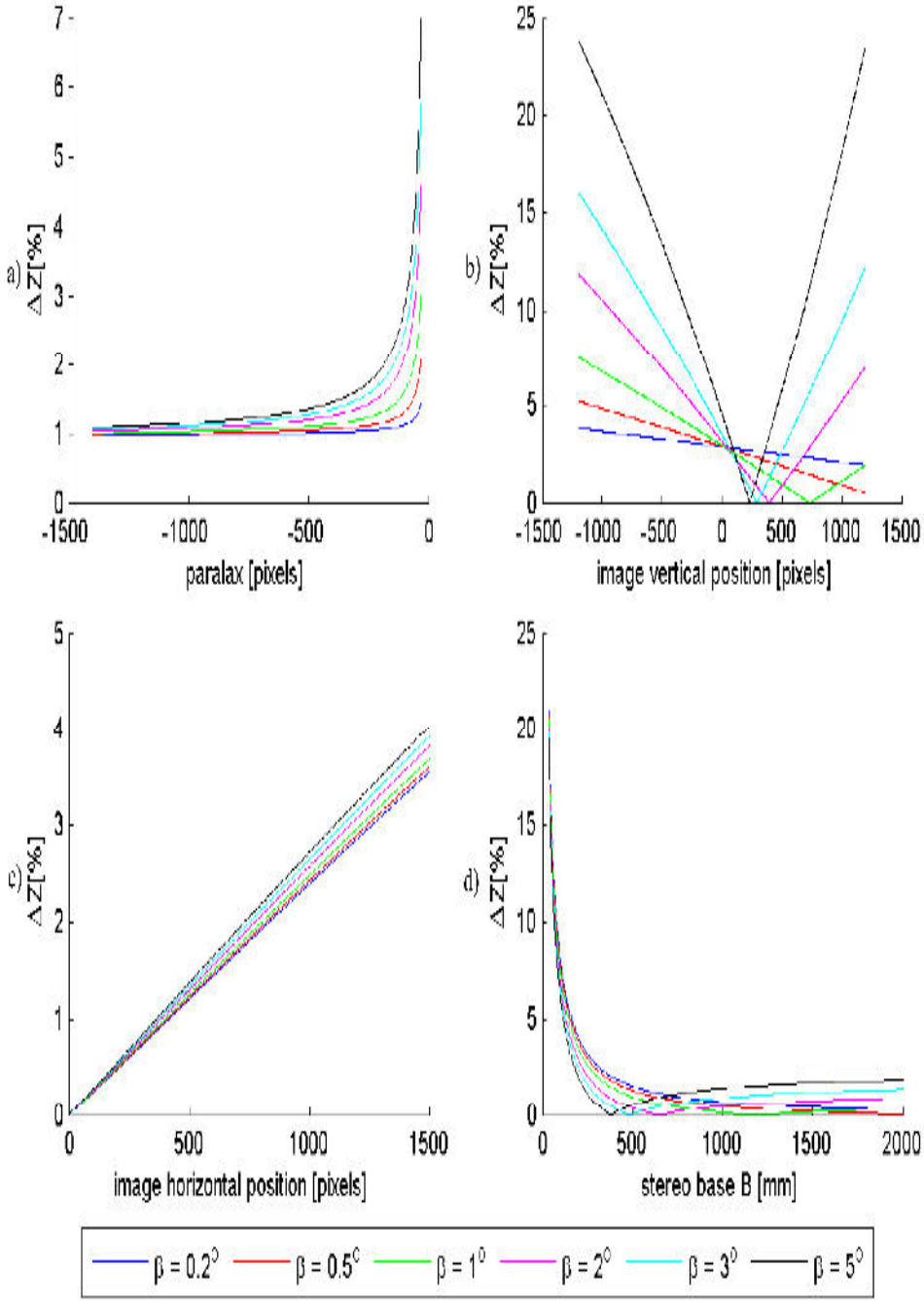


Fig. 3.19: The dependency of the relative error ΔZ in the horizontal space coordinate Z on the a) horizontal parallax b) image vertical position, c) image horizontal position, d) stereo base. The fault angle β is parameter. Used parameters of the camera system B=500mm, f=8.5mm.

Subsequently, we assume that the first camera is perfectly calibrated and its optical axis represents the z axis of the ordinate system with the center in the focus. The calibration of the second camera is perfect except for a certain rotation angle γ about the y axis.

The general formulas for calculating error derived above (3.13), (3.16) and (3.10) can be used again. Therefore, equations for calculating x'_2 and y'_2 have to be derived. The planar model of the situation (see Fig. 3.20) was used in [49]. Subsequently, the formula for calculating the error ΔZ 3.43 was derived and by using the following mathematical procedure and operations we obtain the following formula

$$tan\left(\gamma+\omega\right)=\frac{x'_2}{f_c}, \tag{3.39}$$

$$tan\omega=\frac{x_2}{f_c}. \tag{3.40}$$

Subsequently, using trigonometric relationships $tan\left(\gamma+\omega\right)=\frac{(tan\gamma+tan\omega)}{(1-tan\gamma\cdot tan\omega)}$ we obtain

$$\frac{x_2}{f_c}=\frac{tan\gamma+\frac{x'_2}{f_c}}{1-tan\gamma\cdot\frac{x'_2}{f_c}} \tag{3.41}$$

where γ is the yaw angle between two cameras.Then x'_2 can be expressed by the following equation:

$$x'_{2,Z}=f_c\cdot\frac{x_2-f_c tan\gamma}{x_2 tan\gamma+f_c}, \tag{3.42}$$

We obtain the relation for error by substitution (3.42) into (3.10).

$$\Delta Z\simeq-\frac{tan\left(\gamma Z^2_{real}\right)\left[1+\left(\frac{y_2}{f_c}\right)^2\right]}{B},$$

$$\Delta Z\simeq-\frac{tan\gamma\left(Z^2_{real}+Y^2_2\right)}{B}. \tag{3.43}$$

The experiment for verifying its correctness was executed. The experiment is again based on the position of found points in the rendered images. The experiment revealed that formula (3.42) is in accordance with the real state. However, [49] does not consider error in other spatial coordinates and error in image horizontal coordinate y. A spatial model of this situation was used for deriving these errors. This situation is more complicated for this error. The focus of the camera changes its position. The basic of the derivation is finding the point of intersection of the plane ϑ and line segment $\overline{u}$. Firstly, the line segment $\overline{u}$ passes points $P[X,Y,Z]$ and $F'[0,f_c\sin\gamma,f_c\cos\gamma]$. Then the line segment $\overline{u}$ is described by the parametric equation

$$U=P-tF', \tag{3.44}$$

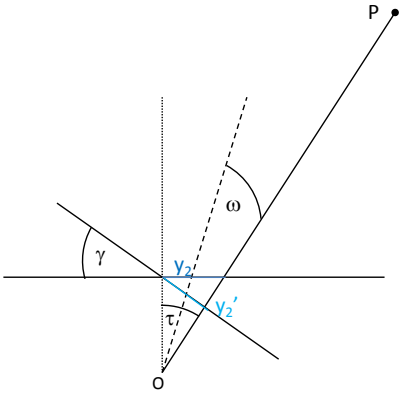


Fig. 3.20: The planar model of the geometric situation for error in yaw (used in article [49]).

The plane ϑ is described by a general equation by using three points which lie on it $P_1[0,0,0]$, $P_2[1,0,0]$ and $P_3[0,\sin\gamma,\cos\gamma]$.

$$0 = ax + by + cz. \tag{3.45}$$

Subsequently, the segment line equation is substituted to the plane general equation. Then, parameter t is computed and substituted back to the (3.45). After this substitution, the final position of x'_2 and y'_2 is obtained. After mathematical operations, the simplified formulas are obtained

$$x'_{2,P} = -\frac{X f_c \cos\gamma^2 - Z f_c \cos\gamma \sin\gamma}{Z\cos\gamma - f_c + X\sin\gamma}, \tag{3.46}$$

$$y'_{2,P} = \frac{Y f_c}{Z\cos\gamma - f_c + X\sin\gamma}. \tag{3.47}$$

Subsequently, the experiment comparing Zhao's and the proposed formula for calculating the change of image point position is performed. The results are in Tab. 3.3. It is obvious from the table that our newly derived formulas (3.46) and (3.47) are usable. Consequently, these derived formulas are successively substituted to formulas (3.10), (3.13) and (3.16). The final equations for errors in all spatial coordinates are obtained

$$\begin{aligned}\Delta X_T = \frac{(B-X)\left(\frac{Z^2\sin 2\gamma}{2} - X f_c + XZ\left(\cos\gamma - \left(\frac{\cos 2\gamma}{2} + 0.5\right)\right)\right)}{f_c(B-X) + \frac{Z^2\sin 2\gamma}{2} + \sin\gamma\,(X^2 - BX) - BZ\cos\gamma} \cdot \ldots \\ \ldots \frac{+X^2\sin\gamma}{+XZ\left(\cos\gamma - \left(\frac{\cos 2\gamma}{2} + 0.5\right)\right)},\end{aligned} \tag{3.48}$$

$$\Delta Y_T = Y - \frac{BYZ)}{Z\left(\frac{Xf_c\left(\left(\frac{\cos\gamma}{2}\right)+0.5\right)-\frac{Zf_c\cos 2\gamma}{2}}{Z\cos\gamma - f_c + X\sin\gamma} + \frac{f_c(B-X)}{z}\right)}, \tag{3.49}$$

$$\Delta Z_T = \frac{Xf_c\left(\frac{\cos\gamma}{2}+0.5\right) - Zf_c\cos\gamma\sin\gamma}{Bf_c\left(Z\cos\gamma - f_c + X\sin\gamma\right)} - \frac{XZ_1}{B}. \tag{3.50}$$

Subsequently, the obtained formulas were verified by the same procedure as for formulas (3.23), (3.24), (3.25). Therefore, the differences between spatial coordinates in an ideal camera stereoscopic system and a system with error in alignment were calculated and compared with the theoretical error obtained by the newly derived formulas. The results for a few points are in Tab. 3.3. Theoretical errors and real differences are equal, therefore the derived formulas are verified.

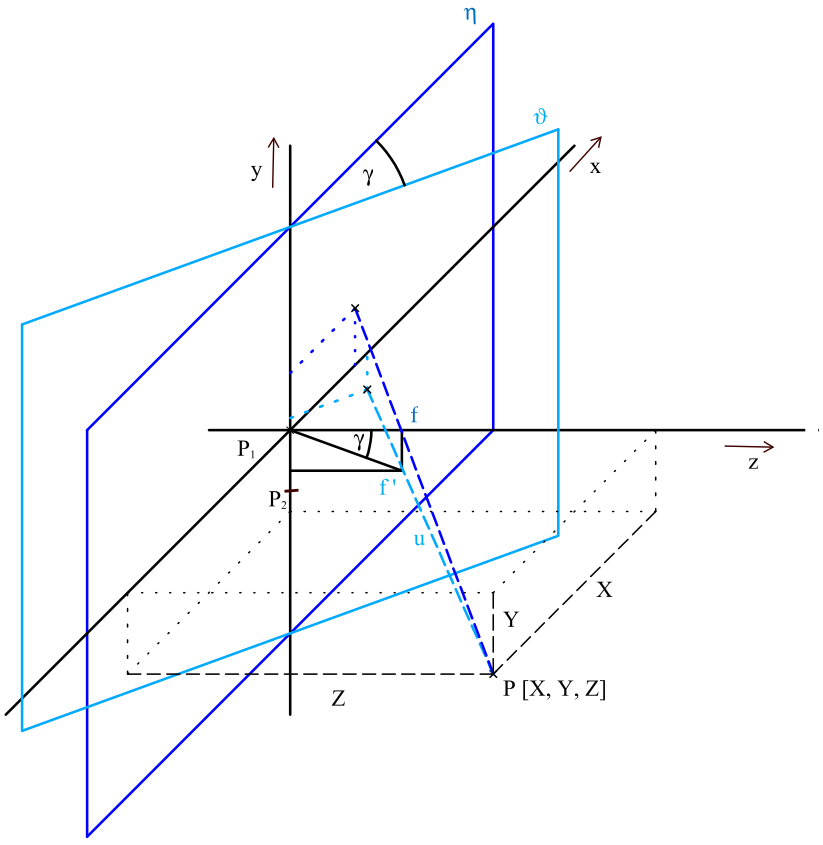


Fig. 3.21: The model of the geometric situation for yaw error. The dark blue plane represents the plane of the image without error. The skyblue plane represents the plane of the image with error. The formulas error of the image coordinates (3.46) and (3.47) are derived from this image.

	pixel 1	pixel 2	pixel 3	pixel 4
x_1 [pixel]	554.00	452.00	427.00	451.00
y_1 [pixel]	300.00	197.00	351.00	300.00
x_2 [pixel]	400.00	298.00	273.00	297.00
y_2 [pixel]	300.00	197.00	351.00	300.00
$x'_{2,F}$ [pixel]	575.00	472.00	448.00	472.00
$x'_{2,Z}$ [pixel]	574.00	472.65	447.50	472.65
$x'_{2,P}$ [pixel]	573.64	472.1	447.35	472.1
$y'_{2,F}$ [pixel]	300.00	197.00	351.00	300.00
$y'_{2,Z}$ [pixel]	—	—	—	—
$y'_{2,P}$ [pixel]	300.00	197.46	349.72	298.66
X_{REAL} [mm]	0.00	198.70	247.40	200.65
X_{ERROR} [mm]	2731.150	1042.90	698.06	1028.60
ΔX_C	2731.30	844.20	450.66	827.95
ΔX_T	2682.02	877.30	450.59	877.30
Y_{REAL} [mm]	0.00	200.65	-99.35	0
Y_{ERROR} [mm]	0.00	1672.05	847.72	0.00
ΔY_C	0	1672.05	847.07	0
ΔY_T	0	1731.15	3749.35	1731.15
Z_{REAL} [mm]	3896.10	3896.10	3896.10	3896.10
Z_{ERROR} [mm]	31579.00	28571.00	29485.88	32467.10
ΔZ_C [mm]	35475.10	32467.10	33382.00	33748.1
ΔZ_T [mm]	34445.10	33748.1	33382.1	32467.1

Tab. 3.3: The verification of the proposed formulas (3.46), (3.46) for calculating error image positions $x'_{2,P}$, $y'_{2,P}$ and formulas (**??**),(3.49),(3.50) for calculation of the error of the spatial coordinates ΔX_T, ΔY_T, ΔZ_T for the yaw of the camera.

Figures 3.22, 3.23, 3.24, illustrate the dependencies of the relative error in space coordinates X, Y, Z on the parallax p_x, vertical x_{im} and horizontal y_{im} image positions and on stereo base B. There is opposite situation than in previous error angle α. Coordinate X is the most sensitive to error in pitch and its relative error reaches a value of about 10 percent for angle 1°. The yaw error is the most critical error from the view of overall error. The parallax p_x is strongly influenced if the image coordinates in one image are strongly changed. The horizontal parallax is in equations for calculating of all spatial coordinates. Subsequently, all three spatial coordinates are critically influenced.

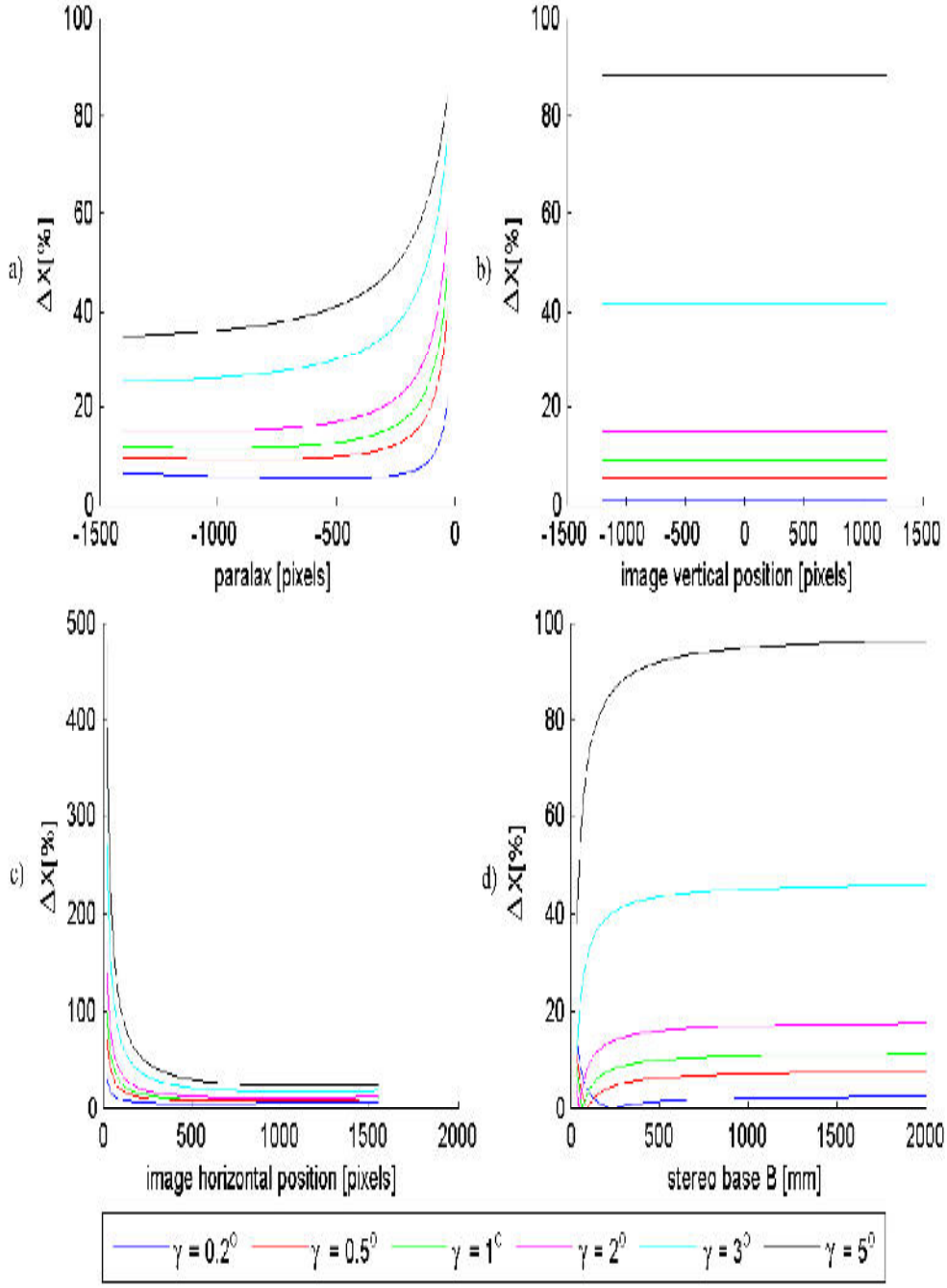


Fig. 3.22: The dependency of the relative error ΔX in the horizontal space cordinate X on the a) horizontal parallax b) image vertical position, c) image horizontal position, d) stereo base. The fault angle γ is a parameter. Used parameters of the camera system B=500mm, f=8.5mm.

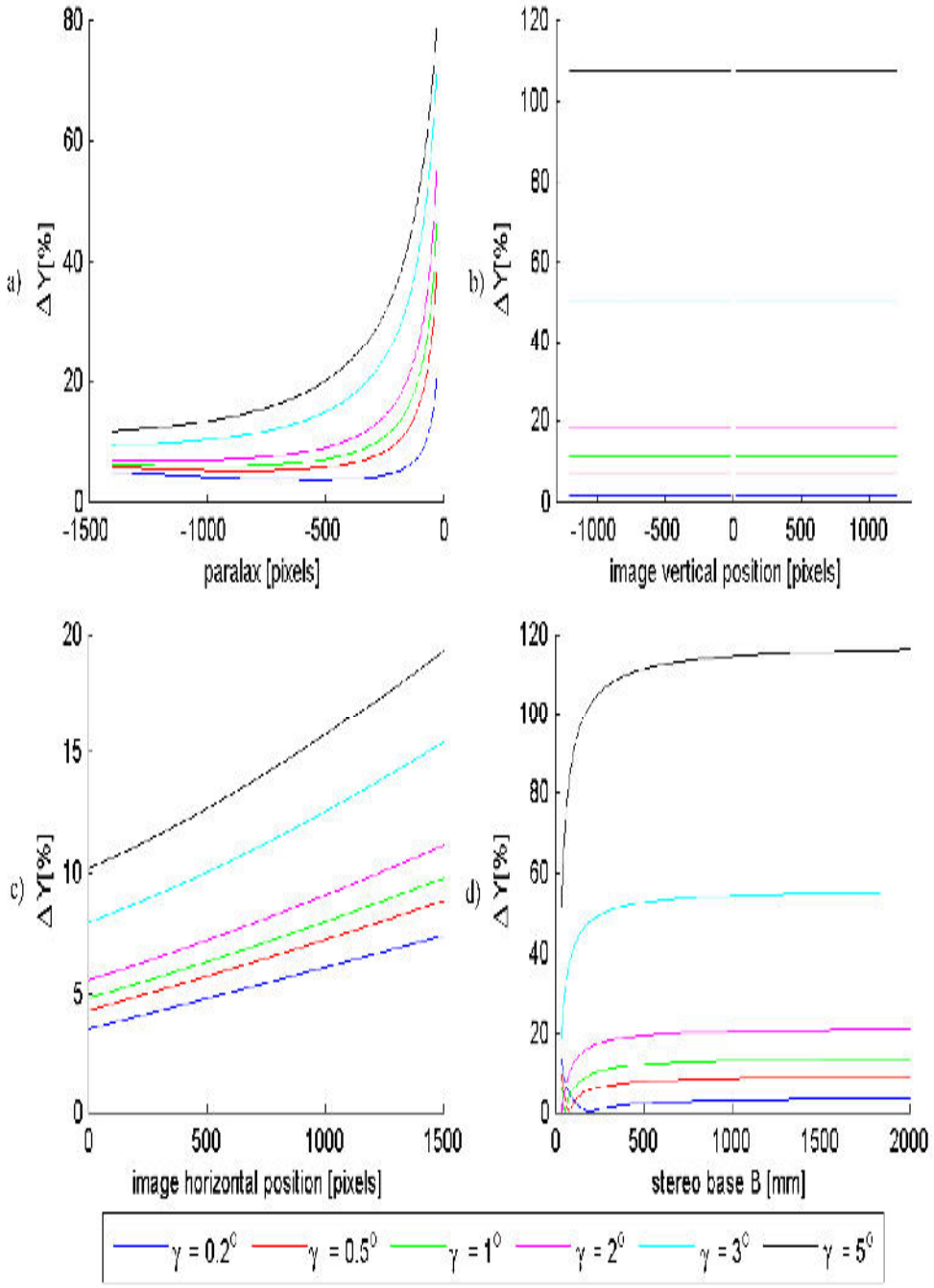


Fig. 3.23: The dependency of the relative error ΔY in the horizontal space coordinate Y on the a) horizontal parallax b) image vertical position, c) image horizontal position, d) stereo base. The fault angle γ is a parameter. Used parameters of the camera system B=500mm, f=8.5mm.

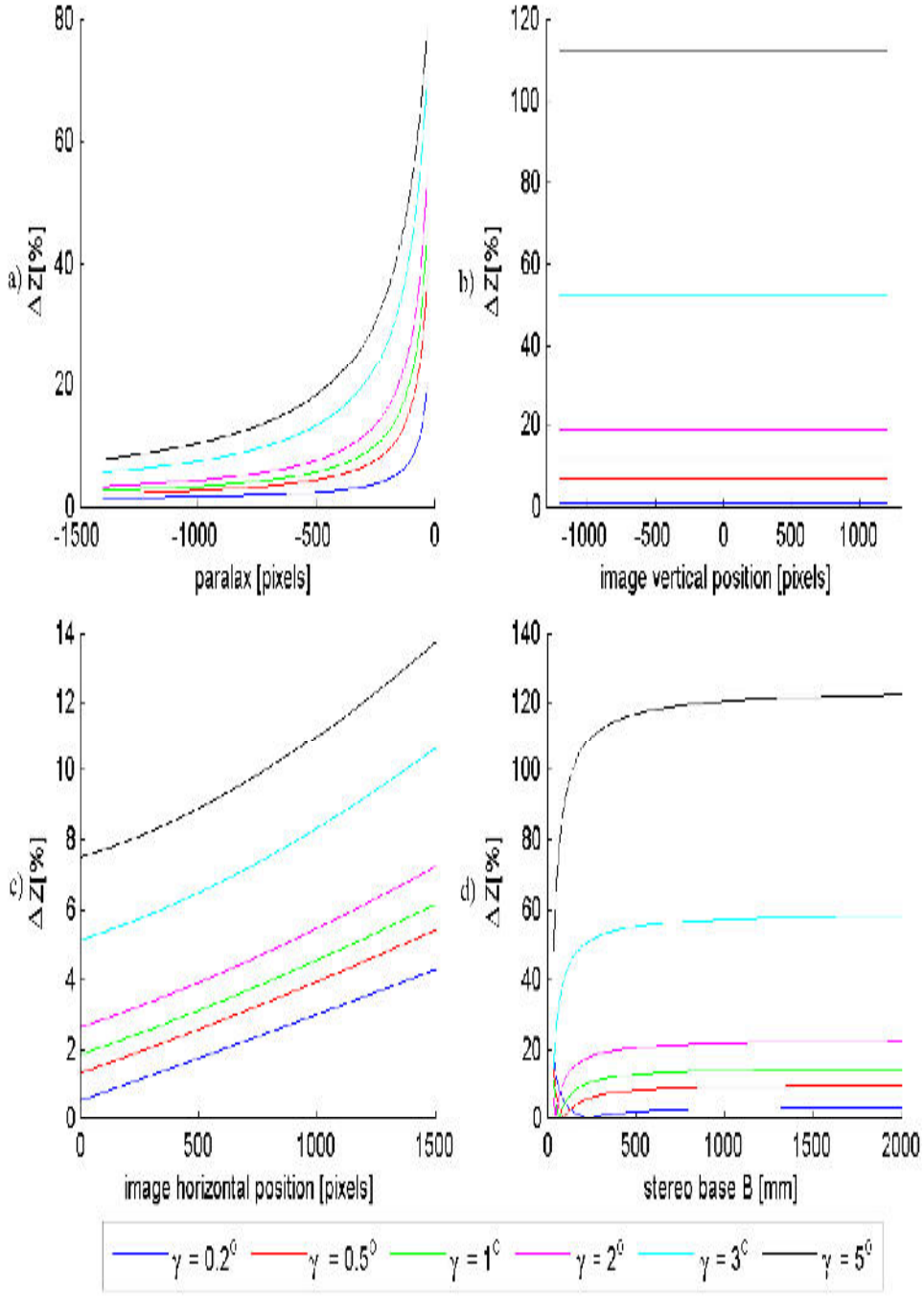


Fig. 3.24: The dependency of the relative error ΔZ in the horizontal space coordinate Z on the a) horizontal parallax b) image vertical position, c) image horizontal position, d) stereo base. The fault angle γ is a parameter. Used parameters of the camera system B=500mm, f=8.5mm.

3.2.2 Errors in general positions of the cameras

The cameras of the 3D sensing system can generally have arbitrary positions in space. Then, the error in camera alignment is equal to the error in rotation matrix **R**, which is obtained by a set of corresponding points. The coordinate system center is usually located at the optical center of the first camera. Therefore, the rotation angles ϕ,κ, ω and matrix **R** represent the relation between both cameras and between the coordinate system and the second camera. Assuming that we know rotation angles ϕ,κ, ω between the optical axis of the camera and axes of the coordinate system, then the theoretical rotation matrix of the camera **R** can be obtained from rotational angle by using relation (3.51) [105]

$$\mathbf{R} = \begin{bmatrix} \cos\phi\cos\kappa & -\cos\omega\sin\kappa + \sin\omega\sin\phi\cos\kappa & \sin\omega\sin\kappa + \cos\omega\sin\phi\cos\kappa \\ \cos\phi\sin\kappa & \cos\omega\cos\kappa + \sin\omega\sin\phi\sin\kappa & -\sin\omega\cos\kappa + \cos\omega\sin\phi\sin\kappa \\ -\sin\phi & \sin\omega\cos\phi & \cos\omega\cos\phi \end{bmatrix} \tag{3.51}$$

On the contrary, the rotation angles can be determined from rotation matrix **R**. The rotation matrix **R** is obtained by using an 8- point algorithm from the set of corresponding points. Therefore, errors in determining corresponding points cause an error in the rotation matrix **R**. The same error in the determining the same corresponding point can variously influence the calculation of the rotation matrix. The errors are influenced by many factors. The resulting error of the rotation matrix is given by the combination of the error in each corresponding point. Therefore, the influence of the error in a particular correspondence to the results is affected by the error in other correspondences. The next aspect is mutual camera positions. The influence of the camera position is obvious from section 3.1, where dependencies illustrate that deviation of the error for an identical set of points can be relatively perceptible if cameras are located in various positions. It is not possible to investigate all possible combinations. All of the following experiments are statistical sensitive analyses. The received results are valid only for specified conditions. However, some general hypothesis and conclusions can be deduced.

In the first experiment, the additive white Gaussian noise with various Signal Noise Ratio (SNR) is added to all accurately found corresponding points. SNR is in the range from 40dB to 60dB. The method Monte Carlo was used. A thousand repetitions of reconstruction with a particular level of noise was executed. Subsequently, the average value, standard deviation and worst case were determined. The experiment can be described in several steps:

- determination of the rotation matrix $R_{accurate}$ by using accurate corresponding points,

- determination of the correct mutual angles between cameras (ϕ, κ, ω) by using $R_{accurate}$ and relationship (3.51),
- degradation of the set of corresponding points positions by adding noise (error in all points defined by SNR),
- calculation of the rotation matrix R_{error} from the set of degraded corresponding points,
- calculation of the rotational angles of the cameras ϕ, κ, ω from matrix R_{error},
- error analysis of rotation angles.

Two scenes were used in the experiment (see Figs. 3.25 [82] and 3.1 [104]). Their models are in Figs. 3.26 and 3.3. The results are shown for various images in Tab. 3.4 and 3.5. The fundamental difference between the analyzed scenes is the number of used correspondences: 13 for Matlab scene and 2675 for Cathedral scene.The error in image positions of corresponding points were in the following ranges for various SNRs:

- 40dB: hundredths and tenths of the pixel, up to 1 pixel,
- 45dB: tenths of the pixel, maximum about 1 pixel,
- 50dB: units of the pixel, maximum about 3 pixels,
- 55dB: units of the pixel, up to 5 pixels,
- 60dB: units of the pixel, up to 10 pixels.

The some conclusions can be deduced from the obtained results. The error of angle ϕ is most sensitive to errors in correspondences. The number of correspondences has crucial importance. It is obvious from comparing tables for each scene that errors significantly increase with decreasing number of correspondences. It is obvious that the error of the rotation matrix is considerable even for a very small error in correspondences. Therefore, the importance of correct correspondences is obvious from the results. The average values are much more smaller than the worst case. An increase in standard deviation for decreasing SNR is axiomatic and expected. The executed experiments and obtained results serve for a few purposes:

- demonstration of the importance of correctly finding corresponding points,
- get an idea of how an error can occur,
- obtaining relation between error of stereo camera alignment and errors in finding corresponding points,
- design process for estimating possible error.

In the second experiment, only one point is debased by an accurately defined error. Other correspondences are accurate. The worst case analysis is executed again. The most sensitive point and most affecting point are found. The most sensitive point is such a point which has the largest errors in dependency on error of other points. The error in all thirteen points is successively simulated and errors in all spatial points are monitored. The points for which the sum of errors in

each situation is largest is determined as the most sensitive point. Parallel to the points which the caused the largest errors in all spatial coordinates of other points (most affecting) are found. These two points are marked in Fig. 3.25. Then, the dependence of the overall error in spatial positions of all points on the error in the most affecting point is plotted in Fig. 3.27. The error in horizontal image position is consider in the range from 0.1 to 10 pixels. The error in depth Z is least sensitive to the error in a particular point. This fact applies for error in all points from the set of corresponding points found in this scene. However, this fact is not valid generally. Subsequently, the dependency of the error of the most sensitive point to the errors in other points is plotted in Fig. 3.28.

Fig. 3.25: The images used in the investigation of error during reconstruction caused by incorrect determination of camera alignment and errors in determining corresponding points. The corresponding points are marked by red marks. The most sensitive point is marked by a blue mark. The most affecting point is marked by a green mark [82].

In the next scenario, the error angles α, β, γ were added to the original angles between cameras ϕ, ω and κ. Consequently, the rotation matrix $\mathbf{R}$ was directly corrupted. This scenario follows from the previous which investigated dependency and sensitivity of the error in rotation matrix $\mathbf{R}$ on the errors in corresponding points. The average error and standard deviation in each spatial coordinate was investigated. 2675 points of the scene were found and used in these experiments. The results indicate that the error in spatial coordinates is most sensitive to the angle error ϕ and that the most sensitive to this error is spatial coordinate X. The results are summarized in Tab. 3.6.

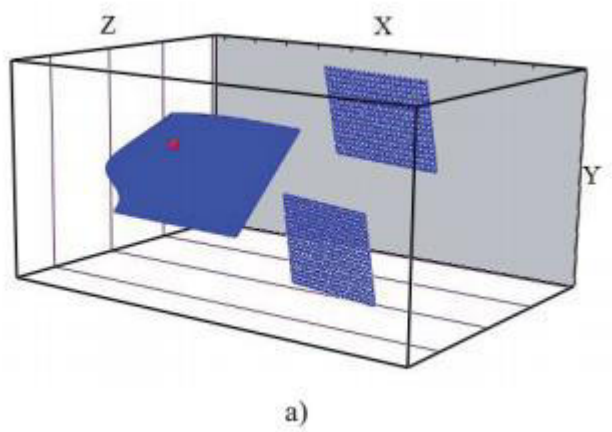


Fig. 3.26: The reconstructed model of scene 3.25 used in experiments. The model is drawn by using 13 reconstructed points.

SNR [dB]	Worst case [°]			Average value [°]			Standard deviation [°]		
	$\Delta\phi$	$\Delta\kappa$	$\Delta\omega$	$\Delta\phi$	$\Delta\kappa$	$\Delta\omega$	$\Delta\phi$	$\Delta\kappa$	$\Delta\omega$
40	35.4	13.7	9.9	12.4	6.2	2.9	8.9	3.4	2.2
45	19.9	9.7	4.8	4.8	3.3	1.1	3.9	2.2	0.9
50	10.1	6.0	2.6	2.2	1.6	0.5	1.7	1.2	0.5
55	4.8	3.7	1.3	1.2	0.9	0.3	0.9	0.7	0.2
60	3.1	2.0	0.7	0.7	0.5	0.2	0.5	0.4	0.1

Tab. 3.4: Results of the Monte Carlo experiment testing the influence of the error in finding corresponding points on the error in rotation matrix for scene in Fig. 3.25.

The other tests were executed for investigation possible situations. However, all results are not presented in the form of plots or tables. The next experiment analyzes the situation with known error in two and more corresponding pairs. The error in one point was compensated by error in another point in some situations. The next conclusion is the fact that error in reconstruction is much more sensitive to one distinctive error in one point than less significant errors in more points. The results prove complexity of the investigated issue. Therefore section 3.2.2 deals with the analysis of possible situations in which the worst case were found, because obtaining the deterministic equation for error calculation is a complex problem exceeding the range of this dissertation.

SNR [dB]	Worst case [°]			Average value [°]			Standard deviation [°]		
	$\Delta\phi$	$\Delta\kappa$	$\Delta\omega$	$\Delta\phi$	$\Delta\kappa$	$\Delta\omega$	$\Delta\phi$	$\Delta\kappa$	$\delta\omega$
40	1.50	0.16	1.20	1.46	0.11	1.05	0.02	0.01	0.04
45	1.44	0.09	0.96	1.41	0.06	0.89	0.01	0.01	0.03
50	1.39	0.06	0.84	1.36	0.04	0.79	0.00	0.01	0.02
55	1.33	0.13	0.73	1.29	0.09	0.67	0.02	0.01	0.02
60	0.48	0.17	0.14	0.35	0.09	0.12	0.04	0.01	0.02

Tab. 3.5: Results of the Monte Carlo experiment testing the influence of the error in finding corresponding points on the error in rotation matrix for scene in Fig. 3.1.

		Average value [%]			Standard deviation [%]		
		ΔX	ΔY	ΔZ	ΔX	ΔY	ΔZ
$\Delta\phi$	1°	25.68	10.60	11.52	13.00	7.81	8.85
	3°	36.13	14.79	16.14	15.50	10.55	11.73
	5°	39.48	16.12	17.67	15.93	11,30	12.63
$\Delta\kappa$	1°	6.86	3.78	6.20	5.15	2.77	4.05
	3°	16.14	6.72	10.97	12.87	7.12	8.66
	5°	29.63	12.08	17.45	58.06	37.61	31.71
$\Delta\omega$	1°	5.11	3.46	5.43	2.39	1.58	2.77
	3°	10.93	4.42	9.29	5.44	3.34	5.90
	5°	15.26	5.87	12.02	7.24	4.50	7.99

Tab. 3.6: Average values of the errors in spatial coordinates depending on the errors in rotation matrix $\mathbf{R}$ for scene shown in Fig. 3.25.

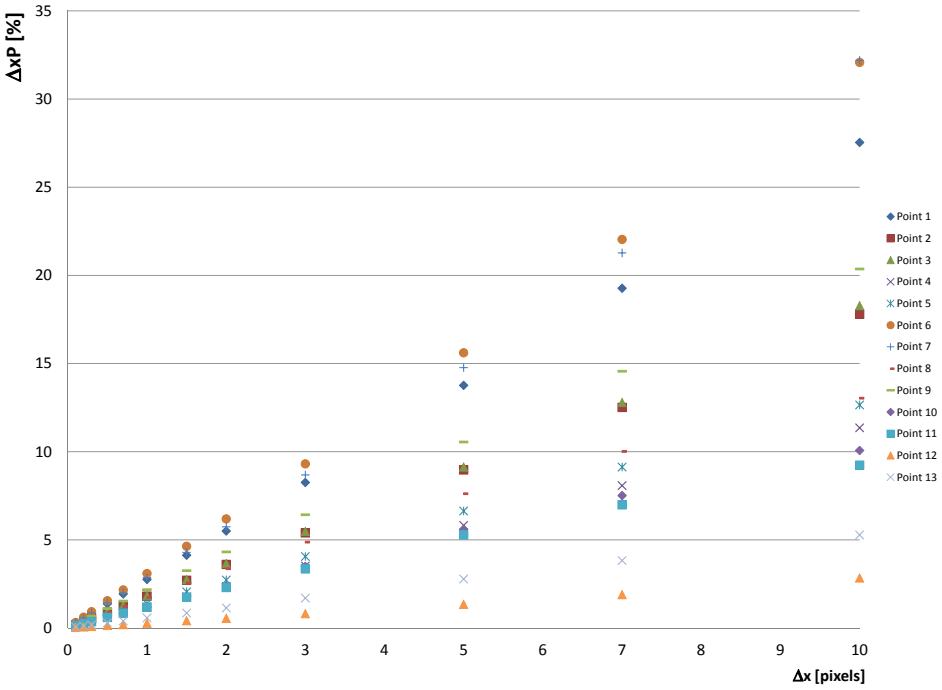


Fig. 3.27: Dependency of the error of spatial position for individual points on the error of horizontal image coordinates x of the most affecting point.

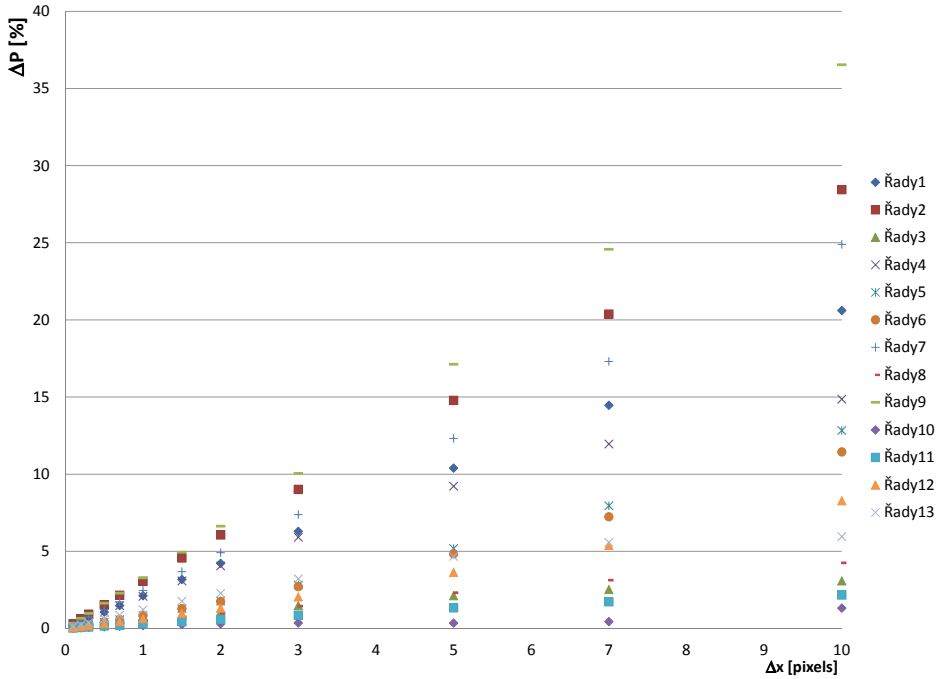


Fig. 3.28: Dependency of the error of the spatial position for the most sensitive point on the error of horizontal image coordinates x of individual points.

4 DEPTH MAP GENERATION

This chapter describes two proposed methods for depth map generation. The first method addresses the passive system for generating the depth map. The proposed system is semiautomatic and it can work without intervention from the user. However, the quality of the resulting depth map can be positively affected by setting some parameters. The system is based on the combination of various approaches. The fundamental ideas use space continuity of the depth map, image segmentation and accurate of finding corresponding points in both images. The method for finding corresponding points proposed in section 2.3 and 2.4 can also be used for this proposed method. The proposed algorithm is implemented in the application described in section 2.5. The second method is based on combining passive and active methods for estimating the depth map. The resulting depth map is obtained as the fusion of the depth maps form each method. The proposed system was created in cooperation with Ing. Kaller. The system includes scanning the scene, stereo sensing and subsequent image processing. I was engaged especially with programming part of this system. In chapter 4.2, the fundamental idea is explained. The algorithms of shadow detection and the combination of both depth maps are described. Finally, some results are presented.

4.1 Algorithm based on similarity measurements and space continuity

Methods for generating a depth map by using stereo matching are proposed in this chapter. The aim of stereo matching is to compute the disparity (mutual spatial shift) of two input images for each pixel. The principle of this approach is in accordance with functions of the human visual system for depth perception. It means that the inputs are two partial images which represent the view of the scene by each eye. Therefore, the images are called left and right. **The images differ only by horizontal parallax**. The horizontal parallax is various for various pixels. Subsequently, the depth of the point in the scene is given by the parallax between points which represent this point in both images; these points are called corresponding points. Therefore, we need to determine these parallaxes (disparities) between the corresponding points. Consequently, we have to find pairs of corresponding points. This issue is frequently based on similarity evaluation of the pixel and finding the best match. One group of methods is based on using legitimate metrics for the similarity of pixels. The basic metrics are SAD – Sum of Absolute Differences, SAS - Sum of Square and Correlation. Instead of points of correspondences, we can

find correspondences of a small areas in an image. These segments can consist of a set of pixels in one row or we can use a segment obtained by some segmentation methods. This approach is used in the algorithm proposed in this chapter. Another approach is based on using sparse point's correspondences. In the first step, significant points in both images are found. A points descriptor or detector are used (for example SUFR, SIFT). Subsequently, the point's correspondences are determined. The sparse point's correspondences can be enhanced in the depth map (dense correspondences) using various methods which are often based on segmentation and dissemination of the information.

The proposed method consists of two fundamental steps. The procedure is described in the flowchart in Fig. 4.1. In the first step, the initial depth map is obtained by implementing SAD (Sum of Absolute Differences) and CGRAD (Cost from Gradient of Absolute Differences) [106]. Briefly, a description of creating the initial map is explained in the following section 4.1.1. The initial disparity map has many discontinuities and errors. Pixels without determined depth occur if metrics for similarity do not have sufficiently reliable disparity for an appropriate pixel. The assignment of depth to these pixels is necessary. Therefore, the next step is required. We proposed an approach for improving the initial depth map. The proposed method utilizes a combination of a some information. The approach is based on the assumption of continuity of the depth map in rows and utilizes information about edges in images. The edge representation of the image has an important role in this process. The proposed method is described in detail in section 4.1.2.

4.1.1 Creation of initial depth map

The process of creating a depth map is the first step in the proposed procedure. This process is based on the algorithm implemented by Shaun Lankton [107]. The algorithm works with an image in three-component representation. The possibility of using an HSV image or pseudo color image was investigated. However, using classic RGB representation with true colors was selected as the most appropriate. The algorithm can be described by the flowchart in Fig.4.2. Input parameters for this process are maximal disparity ($disp_{max}$), win_{length}, $depth_{tolerance}$, and $grad_{weight}$. Their meaning is successively explained. In the first step, the gradient images are obtained which are used for calculating CGRAD. We calculate SAD and CGRAD for a shifted image according to the following equation. The shifting d represents various disparities from minimal to maximal, hence from 0 to $disp_{max}$

$$diffG\,(i,d) = sum|\nabla L_i - \nabla R_{i-d}|, \tag{4.1}$$

$$SAD\,(i,d) = sum|L_i - R_{i-d}|. \tag{4.2}$$

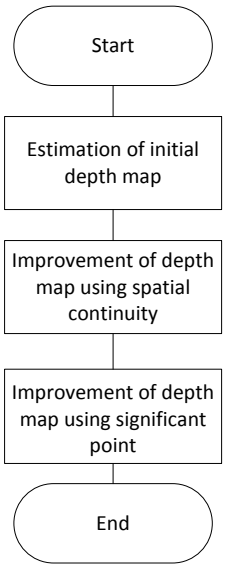


Fig. 4.1: Flowchart of the proposed algorithm for generating the depth map based on similarity measurements and space continuity.

Difference between gradients is calculated for three directions. Therefore, differences are summed

$$CGRAD\,(i,d) = diffG_x + diffG_y + diffG_z. \tag{4.3}$$

Subsequently, the obtained parameters SAD and CGRAD are averaged with window of size Win_{length}.Therefore, the input Win_{length} determines the size of the smoothing filter. Finally SAD and weighted CGRAD are summed

$$difference\,(i,d) = grad_{weight} \cdot CGRAD\,(i,d) + SAD. \tag{4.4}$$

Therefore, the contribution of CGRAD is given by parameter $grad_{weight}$. In the next step, we select a minimal difference for each pixel, and respective disparity is elected as correct for the appropriate pixel. Subsequently, the outputs of this step of the algorithm are two matrices. The first matrix, *depth*, contains disparities for all pixels. The second one, $diferrences$, contains differences for all pixels. This process is carried out in both directions; this means that we found parallax with minimal differences for all pixels in the left and even the right image. Therefore, we have two disparity maps ($depth_{L-R}$ and $depth_{R-L}$). In subsequent steps, we obtain the initial depth map by using the algorithm of type winner take all. This process can be described by the following conditions. If the difference between $depth_{L-R}$ and $depth_{R-L}$ for a given pixel is higher than $Depth_{tolerance}$ then the depth of a

particular pixel in the resulting depth map is set to zero. If the difference between $depth_{L-R}$ and $depth_{R-L}$ for a given pixel is smaller than $Depth_{tolerance}$, then the differences $diferences_{L-R}$ and $diferences_{L-R}$ are compared. The resulting depth is set to $depth_{L-R}$ if $diferences_{L-R} < diferences_{R-L}$. The resulting depth is set to $depth_{L-R}$ for the opposite inequality $diferences_{L-R} > diferences_{R-L}$.

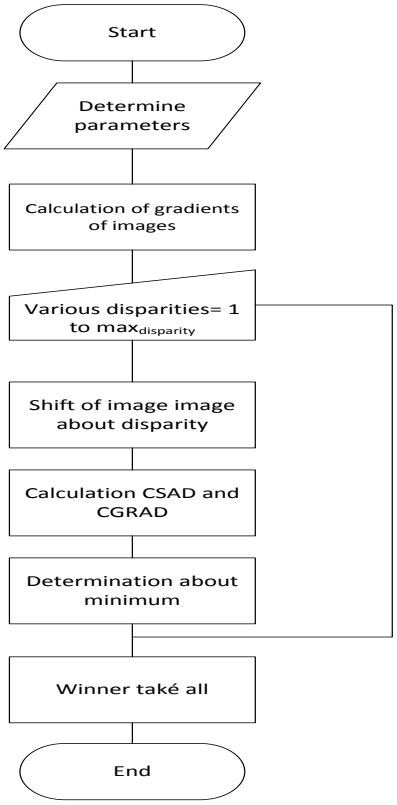


Fig. 4.2: Flowchart of creating the initial depth map.

4.1.2 Improvement of the depth map

In the initial depth map remain areas with undefined depth. Moreover, in the first step, we executed post processing of the depth map. We eliminated small depth regions with big contrast with their surroundings. This is probably due to the depth being determined incorrectly. Subsequently, we eliminated pixels with unreliable depth. In the case of a pixel having unreliable depth, we determine a pixel with a difference (above, in equation 4) exceeding a certain threshold. Elimination means that we set the depth of a particular pixel or region to zero. We want to assign the correct depth to these zero areas. We proposed a solution to this problem. The proposed approach works in individual rows. The solution used the assumption

about space continuity of the depth map. Edge representation of the image is obtained by implementing the Canny detector. The core of the approach follows: The zero regions are found. Subsequently, we find the depth on both boundary of the regions ($depth_{Rborder}$ and $depth_{Lborder}$) and length of the region ($length_{zero}$). Then, delta δD is calculated using the following equation:

$$\delta D = \frac{depth_{Rborder} - depth_{Lborder}}{lenght_{zero}}. \tag{4.5}$$

Parameter δD characterizes rapidity of change of the depth. In the next step we use edge representation. There are four various cases possible. If delta or length is smaller than the threshold, then we use the following equation to calculate depth:

$$depth = depth_{Lborder} + \delta Di. \tag{4.6}$$

Where i is the order of the pixel in the zero region.

In other situations, we use edge representation. Depending on the presence of the edge on the boundary of the zero regions, we use one of three possible abrupt changes of the depth map. All various scenarios are shown in Fig. 4.4 .

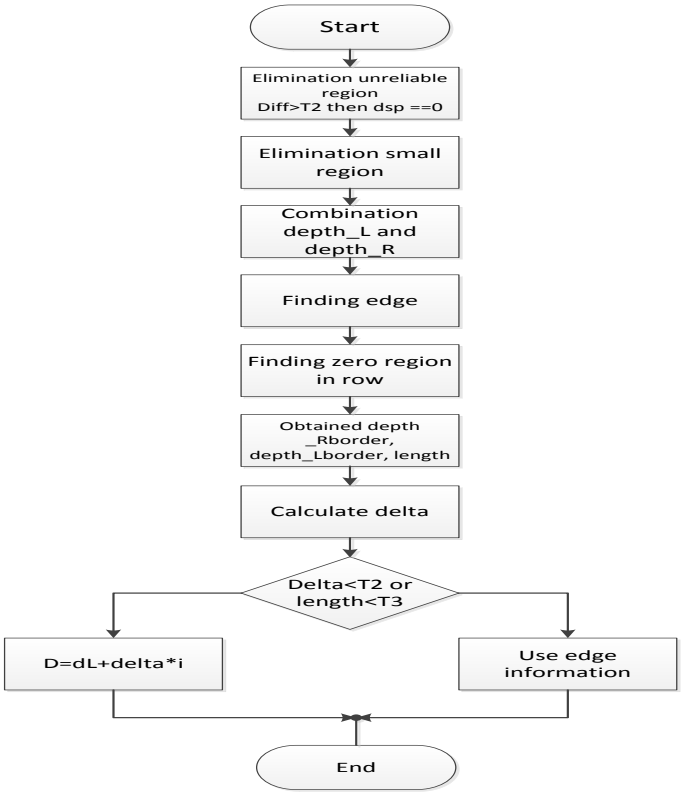


Fig. 4.3: Flowchart of improving depth the map based on space continuity.

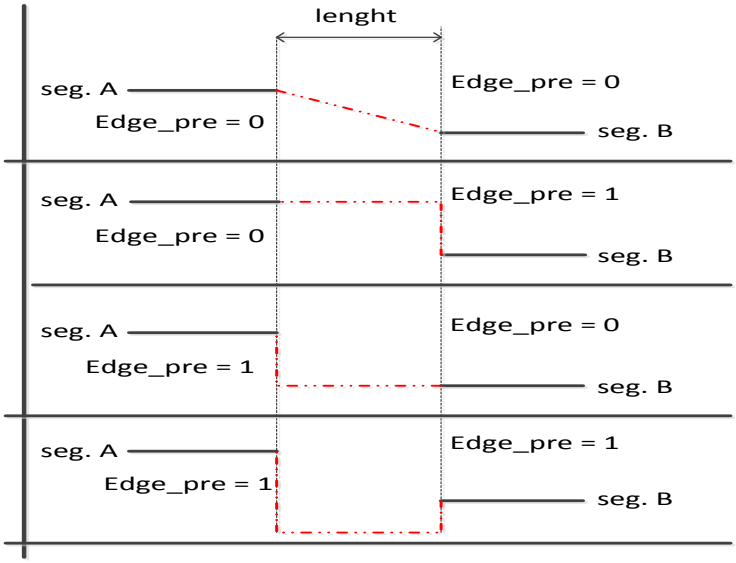


Fig. 4.4: Diagram of the four possible alternatives in the process using edges. A and B are two segments with well determined depth. The zero segment lies between them. The resulting depth is depicted by a red line.

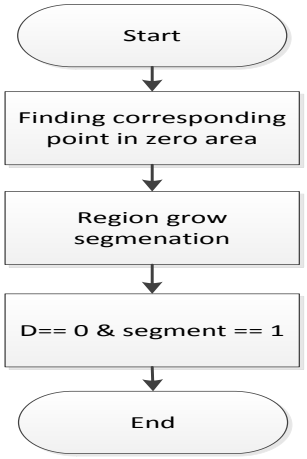


Fig. 4.5: Flowchart of the process to improve the depth map using significant points.

In the second proposed approach, we used segmentation and found the corresponding points. In this case, the algorithm works in two dimensional space. At first, we again find the zero area. Subsequently, significant points are found by the algorithm SURF (see in section 2.2.3). We used SURF with these parameters: Hessian threshold = 0.0001, octaves =5, sampling step in image = 2 , bits of the descriptor = 64. Each significant point is described by fifty sixty-four bit numbers in a range from zero to one. This description is used to finding corresponding points in both images. Points with the smallest difference between descriptors are denoted as corresponding points. In the next step, we detect which significant points from the

found set belong to the individual zero area. We only keep points in the set which lie in the zero area. Subsequently, we executed segmentation using the method "growth from seed". In our case, all seeds are significant points. Subsequently, disparity belonging to the appropriate significant point is assigned to the whole found segment. In this application, two issues are very important. The reliability of finding corresponding points plays an important role. This task is simpler due to the fact that we consider stereo images with corresponding points in the same row. Over segmenting is important too, because we want to prevent gaining too big segments with various depths.

4.1.3 Experiment and results

We implemented the proposed method in MATLAB. The method has a working designation of Depth Continuity Method DCM. Subsequently, we performed some tests of the applicability of the method. For this purpose, we used images from the open database Middlebury Stereo Datasets [89]. This database contains stereo images and a true depth map of the scene. The used images have a size of 370x465 pixels. The obtained results are compared to results obtained by other ways:

- CSAD+belief propagation BP,
- commercial software Stereo tracer ST.

The obtained depth maps were compared with true depth maps which are part of the used database [89]. The reliability is given as the ratio between the number of pixels with correctly determined depth and total number of pixels. The average error in an individual image is calculated as the average difference between depth of the particular pixel estimated by the respective method and the depth in true depth map. The results are summarized in Tab. 4.1. Some resulting depth maps and input images are shown in Fig. 4.6.

image	Reliability [%]			Average error in depth [pixel]		
	ST	BP	DCM	ST	BP	DCM
Tsukuba	34.8	60.8	79.4	28.1	10.0	8.6
Rock2	40.2	35.1	91.9	39.9	43.8	2.2
Baby1	43.8	79.9	86.2	16.4	20.1	8.4
Cloth3	41.0	49.2	94.6	25.9	22.3	2.76

Tab. 4.1: The reliability and average error of the depth map estimated by various methods.

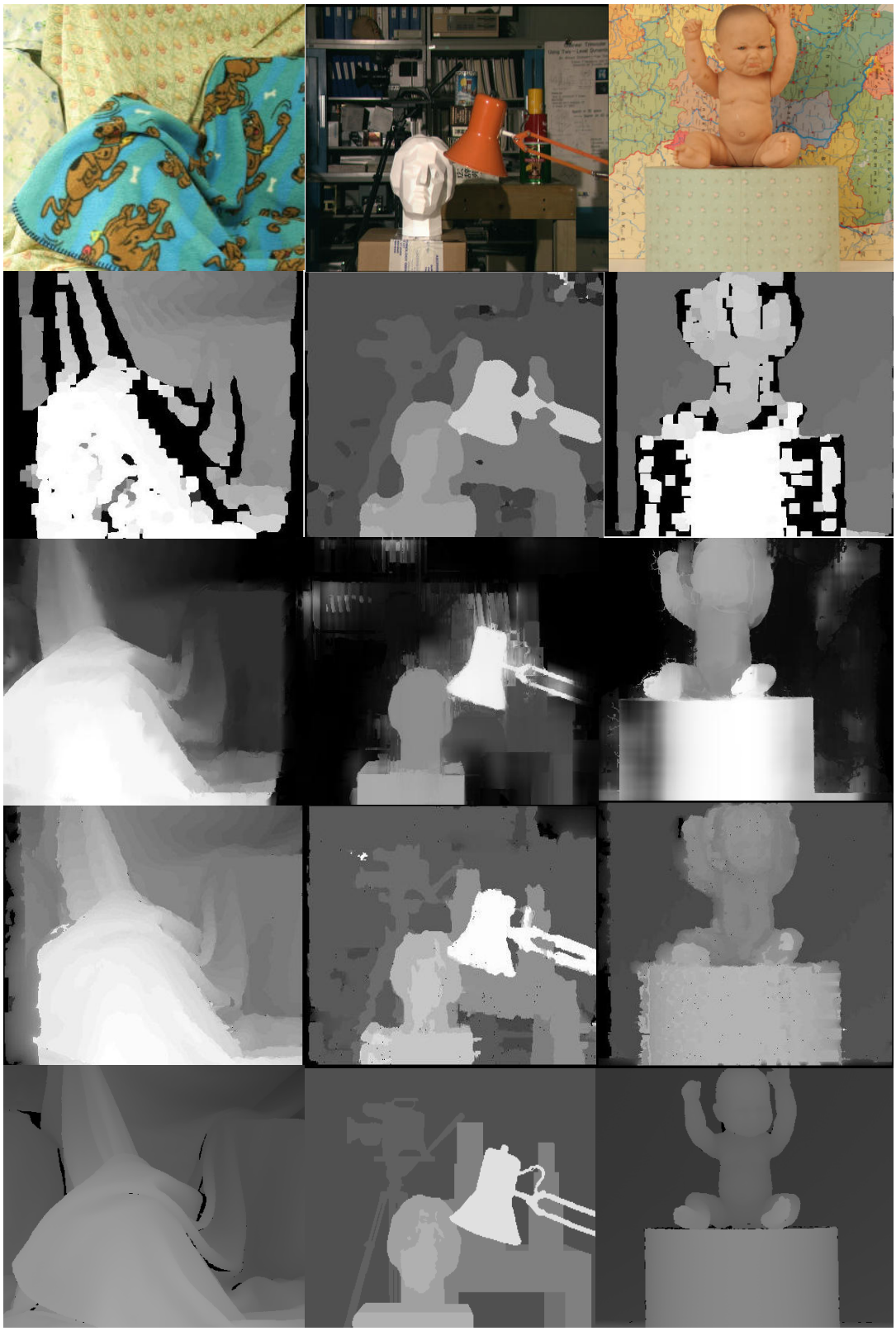


Fig. 4.6: Example of the resulting depth map. First row: left input image, second row: the result from the stereo tracer, third row: the result from belief propagation, forth row: the result from our proposed method, fifth row: true depth map.

4.2 Accurate depth map using combination of the passive and active methods

The utilization of combining passive and active approaches is another way of how to acquire a more accurate estimation of the depth map. I cooperated on this topic with ing. Kaller. The active method is incoherent profilometric scanning. Incoherent profilometric scanning [108], [109] is based on the projection of the fringe pattern on a scene. The passive method used the two stereo images [110]. The fundamental idea is utilizing the good properties of both component methods. On the contrary, disadvantageous features of the used method will eliminated by using in this combination. The advantage of incoherent profilometric scanning is a continuous and accurate depth profile for individual objects in a scene. On the contrary, disadvantages of the active method is its failure to maintain the relation between depth of the individual objects in a scene. On the other hand, we can easily obtain distance between individual objects in a scene by using the stereo method. However, the disadvantage of the passive method is its inaccurate profile of the individual object. Discontinuities occur in depth of the individual objects. This error is caused by the matching problem. The depth map from the stereo images is obtained by finding corresponding points and corresponding areas in the left and right image.

The schematic plan of the workplace for obtaining the depth map by combining the passive and active methods is in Fig. 4.7. The workplace contains a DLP projector which projected a fringe pattern on a scene. The projection is controlled by a simple application running on a PC. The scene is captured by a stereo camera. Using a stereo camera accurately captures the required stereo images. The optical axes of the projector and camera cannot be parallel. The last part of the workplace (besides scanned scene) is a reference plane. The reference plane is realized by some flat smooth white board. Subsequently, the image processing part of the method can be executed when we have captured the required images.

Fig. 4.9 shows a flowchart of the image processing. In the first step, the depth maps is obtained depth map from the profilometric and stereo images. The combination of both depth maps is the last step in our method for generating the depth map. The proposed procedure is based on objects detection and subsequently finding the range of the depth for each object in the depth map obtained by the passive method (*stereo_depth_map*). The found range is used for transforming the depth map obtained by the active method (*profilometric_depth_map*). The procedure is deduced from advantageous properties of both depth maps. The procedure is described in more detail in section 4.2.4.

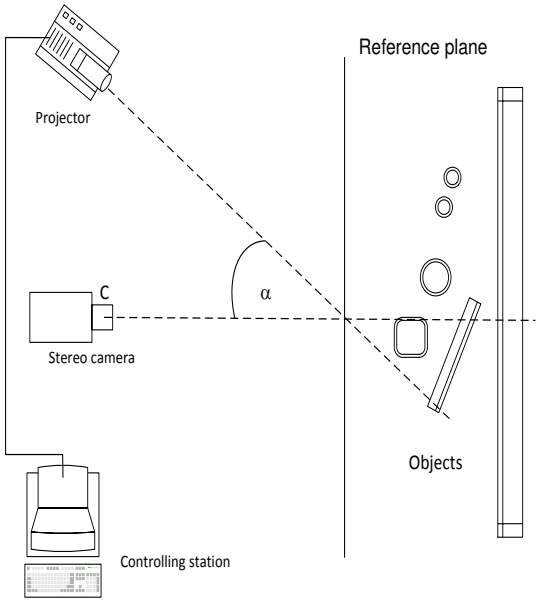


Fig. 4.7: Schematic plan of workplace for combinaing passive and active sensing.

4.2.1 Depth map from stereo image

Most of today's 3D sensing and capturing systems use, for depth map estimation, one of the passive methods based on using two or more images of the analyzed scene. There are mainly two types of these methods. The fundamental difference between the methods is in the cameras' positions. In the first case, the cameras are in general positions (see Fig a) and their optical axes are not parallel. In the second case, the cameras are in the so-called normal positions. The positions of the optical center of the cameras differ only in horizontal direction and their optical axes are parallel. The described distribution has a crucial effect on the usable algorithms. The normal case is more frequently used in applications which were considered. Therefore, the normal position is assumed in our method. This case is simpler because corresponding points are in the same row. This fact brings important constraints to finding corresponding points. Many methods based on the various principles exist for estimating the depth from stereo images. The efficiency of the many algorithms is tested on the webpage of a research team from Middlebury. We used the commercial program StereoTracer [111] in the first tests. Subsequently, we utilized the method from Shawn Lankton [107]. In the final version of the method, we implemented the original proposed method described in section 4.1.

Discontinuities and errors arise in the depth map, obtained by methods based on stereo vision when objects of the scene have a large monochromatic surface or recurring texture, because significant points cannot be identified. Therefore, corre-

spondences cannot be determined. The depth map obtained by the passive method is shown in Fig. 4.12b.

4.2.2 Fringe pattern profilometry

Profilometry is a very commonly used method for accurate surface topography measurement. The coherent light can be used, but in macroscopic scanning systems is usually used incoherent methods using projection of the some pattern by DLP projector. Fringe Pattern Profilometry (FPP) is one of the possible approaches. This approach can be practically used, for example, in biometric identification [114] or in industry quality control [115],[116] ,[117]. Each row of the pattern is a sinusoidal signal. The signal is phase modulated by incidence on the surface of the objects at different distances. Therefore, the three- dimensional profilometry can be obtained by determining the phase difference between the original and deformed pattern. In literature, various methods for converting change of phase to depth, for example: spatial phase detection, Fourier transform profilometry and phase shifting profilometry. In our method we implemented phase shifting profilometry (PSP) which is very easy to implement. The fundamental idea of this method is phase-shifting of the pattern in time. In PSP, N (N $\geq$ 3) shifts of the phase are executed and N frames for projection are formed. The phase shift between the signal in individual frames is 2π/N (see Fig.). Subsequently, the formed patterns are sequentially projected into the reference plane and surface of the measured object, and captured by a CCD camera. We use four shifts in our implementation. Subsequently, the final change of phase is calculated using the following equation [112]

$$wrapped\_phase = arctan\left(\frac{(S_1 - S_3)\cdot(P_2 - P_4) - (S_2 - S_4)\cdot(P_1 - P_3)}{(S_1 - S_3)\cdot(P_1 - P_3) + (S_2 - S_4)\cdot(P_2 - P_4)}\right), \quad (4.7)$$

where $S_{1,2,3,4}$ denote images with projection fringe pattern on the scene (with objects), $P_{1,2,3,4}$ denote images with fringe pattern projected on the reference plane (without objects).

The output of equation 4.7 is in the range $\langle -2\pi, 2\pi\rangle$. The wrapped phase contains sudden changes (wraps) between edge values. We need to eliminate these wraps. For this purpose, unwrapping is executed [118] by implementing the open source code of the method published in [119]. During the experiment, a problem with the shadows has to be solved. The algorithm for unwrapping fail in image areas with shadows. In this area, sudden changes of phase occur frequently and information about phases is lost. The solution is to detect and eliminate the shadows. Shadow detection is a frequent problem. Therefore, many methods for shadow elimination have been proposed, e.g, [120], [121],[122],[123]. A survey of the various approaches for shadow detection is described in an article by Sanin [124].The

algorithm for shadow detection was proposed in [134]. The algorithm is described in detail in section 4.2.3.

4.2.3 Shadow detection in profilometric images

Shadow detection has some specific properties in profilometric images. Due to these properties we can propose specific algorithms. Besides the original picture (further called Object) of the scene, we additionally have a picture with a projected pattern (further called Object_pattern) and a picture with a pattern projected to the background (further called Pattern). Another important fact is that we have a depth map created by a stereo method (further called Depth_stereo). We would like to improve this map by using profilometry. Most methods detect shadows in video and have a more consecutive picture. Therefore, these methods find differences between consecutive images or compare averaged picture with actual image, while our method is based on the fact that shadow regions have low brightness and low contrast.

The proposed method for shadow detection is based on converting the image from RGB to L*a*b. The flowchart is shown in Fig. 4.9. Compared to the previous method this method employs only two images. The first of them is the depth map (further called Stereo_depth_map) and the second one is the original picture of the scene (further called Object). At the beginning we perform smoothing on Depth_stereo. For this purpose, we apply filtration by a lowpass filter in spatial domain. Simultaneously, we convert Object from RGB to L*a*b. Then we work only with the 'a' component, which is suitable for estimating the shadow by thresholding. Consequently, both images are thresholded. All the pixels in the image Object which exceed the threshold Th_object are marked as suspect of belonging to foreground (set equal to 1). All other pixels are set equal to 0. Similarly, all the pixels in the image Depth_stereo which fall within a certain range (defined by Th_min and Th_max) are marked as suspect of belonging to a shadow (set equal to 1) . All other pixels are set equal to 0. The output of the shadow detection is shown in Fig. 4.2.4. The input image is in Fig. 4.11. This image is subsequently used for demonstrating the function of the whole proposed approach. As for the next step, information from both images is combined. The basic assumption says that a pixel cannot be included in the foreground and shadow simultaneously. We combine this hypothesis with the fact that we have a set of points supposedly belonging to shadows (gain using color space L*a*b). This idea is expressed by the following pseudocode

$$
\begin{aligned}
if(S\_S == \&S\_O - 1 \\
Shadow = 1 \\
else \\
Shadow = 0 \\
end.
\end{aligned}
$$

In the final phase, small disturbing artifacts are removed by morphological operations and the MATLAB function bwreaopen.

4.2.4 Combining of the component depth maps

The last step in the proposed procedure is combining the two obtained depth maps. The combination of the depth maps is a very important issue in our method. Inputs to this algorithm are the depth map achieved by the stereo image method, the depth map obtained by the phase shifting method, the shadow map and the original image of the scene. In the shadow map, if a pixel belongs to a shadow, its value is logic 1, else its value is logic 0. The flowchart describing this algorithm is in Fig. 4.10.

The process of combining the depth maps is based on the properties of each depth map. We know, that the stereo depth map provides good information about the mutual position of objects, but the profile of each object is inaccurate. On the contrary, the profile depth map has a accurate profile of each object, however, this method does not provide the relation between positions of the objects. Therefore we want to obtain the profile of each object from the profile map and transform it to the range given by the stereo map.

Firstly, we need to find individual objects in the image. For this purpose, we will use the shadow map and the profile depth map. The next step is based on the assumption that the objects belongs to the foreground, hence the value of the depth map will be high. Concurrently, we assume that objects do not stay in a shadow. In consequence, we will use the following condition. The pixel which satisfies this condition belongs to the object. And its value in the new matrix Object is logic 1.

$$
Shadow\_map == 0 \quad \& \quad stereo\_depth\_map > threshhold\_depthmap. \tag{4.8}
$$

In the following step, we classify objects. The classification of an image means that we define a linking pixel as an object. Output of this step is matrix Class_objects. In the next step, we will find the range of depth of each object. We sort all pixels belonging to the object according to their depth. Subsequently, we determine the upper and lower threshold (th_low, th_up) like values corresponding to 90 and 10 percent of depth of the object. By this way, we obtain the range of depth of each object in the stereo depth map. This range is use as the range of depth of the object

in the final depth map. We find the minimal and maximal depth of each object in profile map (max, min). We transform the depth map by using the above-mentioned parameters of the input depth map and the following relation

$$Result = (th\_up - th\_low) \cdot \frac{depth\_map - min}{max - min}. \tag{4.9}$$

This equation is applied for each object separately and parameters are various for various objects. The final depth map is shown in Fig. 4.12.

Generally, combining various methods and creating hybrid methods brings improvement to the resulting depth map. Therefore this way is perspective. I proposed combining the fringe pattern profilometry and the stereo vision approach. The contribution is designing the procedure. The proposed procedure also used, besides known principles (FPP, unwrapping, stereo visions), two proposed algorithms for executing the component task: shadow detection and synthesis of the two depth maps. The proposed algorithms are described in sections 4.2.4 and 4.2.3. The final depth maps obtained by using the proposed methods are in the Fig. 4.12. The results of this work were published in [78].

Future work in this area will be focused on a hybrid system projecting extra information only on problematic areas where it is impossible to found significant points. Coherent and even incoherent light can serve as extra information. The procedure will be executed in two steps. In the first step, a scene will be captured by a stereo camera and problematic areas will be found. Subsequently, the controlling unit designeded suitable pattern for projection and directs the projector to the problematic area. The system would be semiautomatic in an ideal case. This view is in accordance with other parts of this dissertation. The fundamental aim is obtaining reliable spatial information even about problematic points and areas.

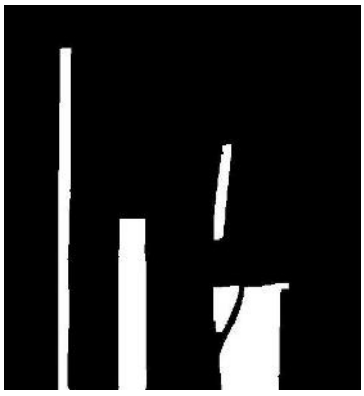


Fig. 4.8: The shadow detected by proposed algorithm in image used in experiment (see Fig. 4.11).

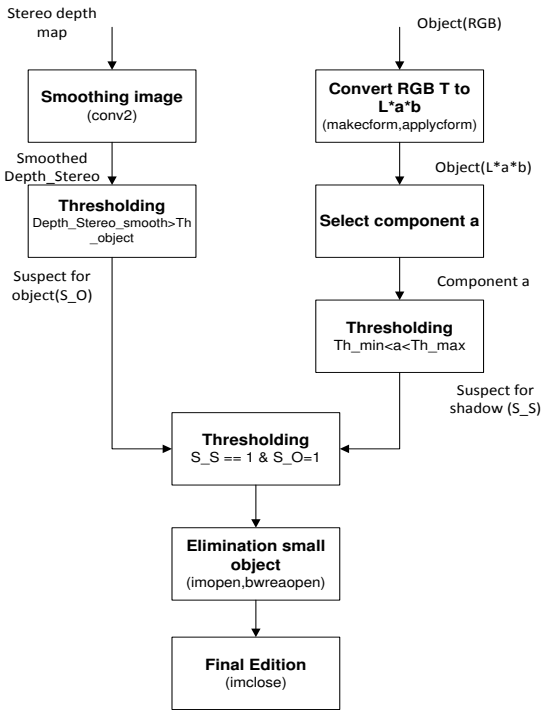


Fig. 4.9: Flowchart of the proposed algorithm for shadow detection based on converting to L*a*b and thresholding.

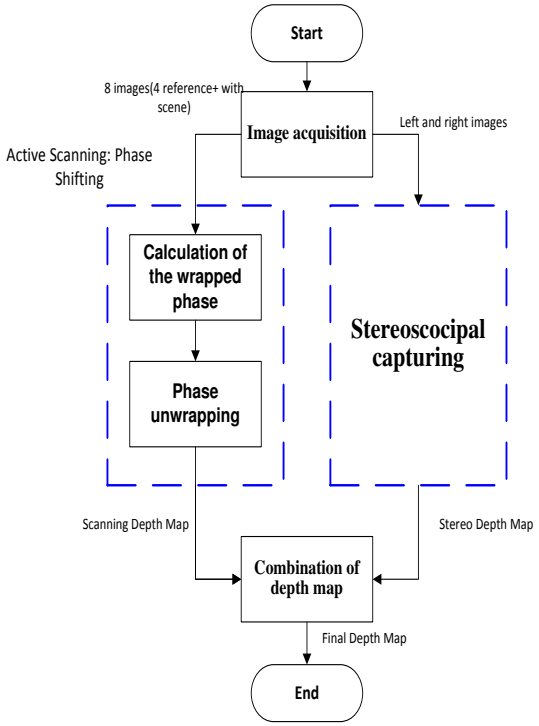


Fig. 4.10: The flowchart of the process of combining the active and passive methods for estimating the depth map.

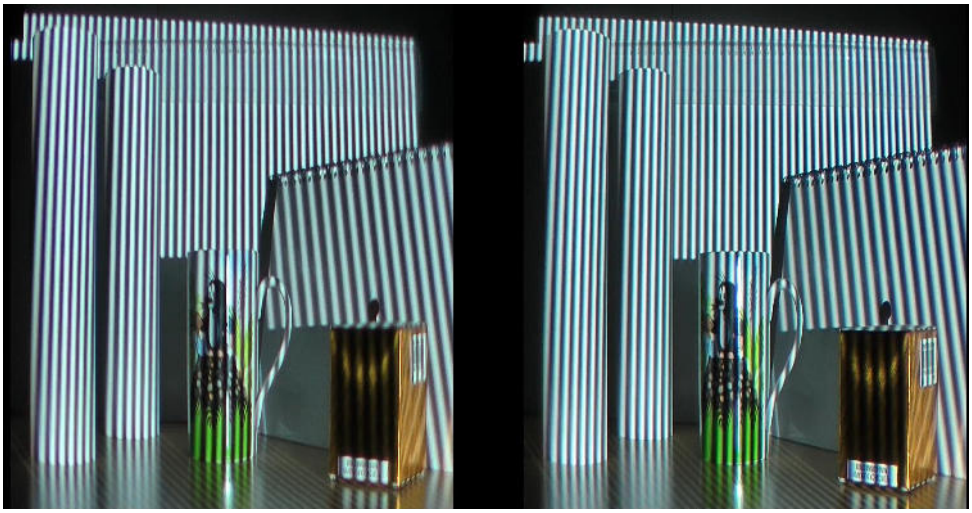


Fig. 4.11: The input image of the scene with projected pattern.

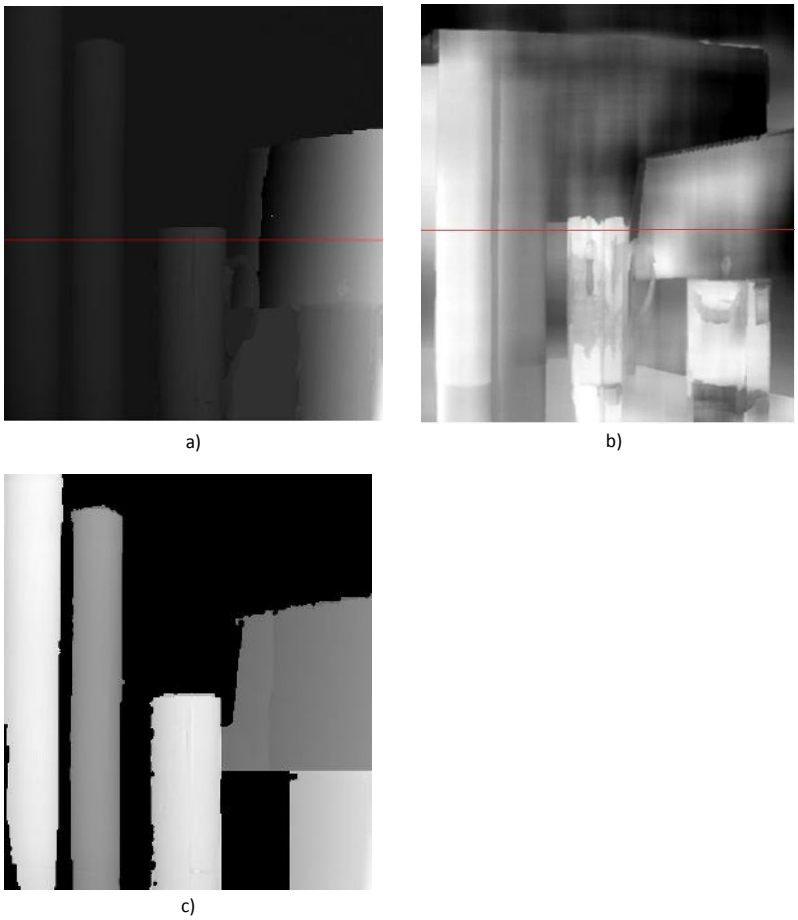


Fig. 4.12: a) The depth map obtained by profilometry. b) The depth map obtained by stereo vision c) The resulting depth map.

5 QUALITY OF EXPERIENCE IN 3D

This section deals with aspects affecting Quality Of Experience (QoE) in 3D video. The topic of 3D video is closely connected with the depth map estimation and its accuracy. 3D television systems have become popular and different systems for 3D imaging are used today. In consequence, a lot of research is devoted to this topic. The important part of this topic is QoE. At the beginning, we have to define the concept of QoE. For a long time, a good definition accepted by majority did not exist. It is a complicated problem. However, one proper definition appeared in White Paper [57] which was created by the consortium Qualinet. This definition says:

Quality of Experience (QoE)**is the degree of delight or annoyance of the user of an application or service. It results from the fulfillment of his or her expectations with respect to the utility and/or enjoyment of the application or service in the light of the user's personality and current state.**

The determination of QoE is a difficult problem, because QoE is affeced by various aspects. The aim of many researchers is to create objective metrics for determining QoE. Subjective tests are a powerfull tool used for this purpose.

5.1 Invitation to evaluating 3D video factors influencing spatial perception

Subjective spatial perception of a 3D image is influenced by many objective and subjective factors. Among the key ones include:

- viewing conditions (viewing angle, room illumination, etc.),
- content of the sequence (parameters of the sequences, spatial activity, range of depth, etc.),
- sensing system,
- 3D imaging system and its parameters (including the quality of video-processing),
- technology and technical parameters of the display (native resolution, frame rate, in the case of LCD the type of backlight too, etc.),
- observer's physiological and psychological features (quality of binocular vision etc.) and others.

During subjective evaluation, these effects cannot be separated. I specified successively the area of my interest. Firstly, I was a member of the research team which organized large subjective tests with three various displaying systems (an active system, passive system and active system used a projector). The main aim of the test

was to compare of three types of display systems. Moreover, we examined a few of the various aspects mentioned above. We investigated the influence of the position of the respondent, hence their viewing angle. Three participants of the test observed one TV display. One of them was in an ideal position in front of the TV. It means that he observed the TV with zero viewing angles in the vertical and even horizontal direction. The other two observers were misaligned in one of the directions (horizontal or vertical). It means that they observed the TV with non zero horizontal or vertical viewing angle. In the aim to cover a variety of different source formats, we used four different sources of video sequences throughout the test. Moreover, we used eight different scenes (sequences) in each source. Roomy illumination was the last investigated aspect. The participants answered six questions. The scale of the evaluation was discrete with seven levels.

- How intense is the 3D effect?
- Judge the depth of the scene?
- Did you feel like you are a part of the scene?
- Did you notice impairments /artifacts in the scene?
- What is the sharpness of the scene?
- Did you experience any uncomfortable feelings?
- Did you feel disturbed by ambient light?

From the results it was obvious that the content of the video sequences have a strong impact on the resulting spatial effect. The experiment proved that all the display technologies under study is comparable in terms of observed intensity of the 3D effect. In an ideal position, the evaluation of the examined quality parameters was without significant differences for different TV systems. However, based on results of the test, I found that it is necessary to examine the dependency of the spatial effect on the viewing angles. Therefore, research focusing on the viewing angle was executed and the results are described in section 5.2.

5.2 Test dependency of QoE on the viewing angle

We made subjective tests of this dependence at the Institute of Radioelectronics Faculty of Electroengineering and Communication Brno University of Technology in 2011. The aim of these experiments was to compare and evaluate the directional dependency of the viewer's spatial perception and 3D image quality on three of today's 3D TVs.

Tests were performed separately on three types of TV displays with different technologies (LCD, plasma) and different 3D displaying methods.. Currently, the most commonly used 3D imaging methods are the following:

- imaging using fade (Eclipse Method). Partial images L, R (left and right) are displayed in the form step by step. The viewer is watching through synchronously driven active glasses. They periodically open a peephole always for a specific data eye for which it momentarily displays an image. This method is particularly suitable for plasma TVs with a short response time, which allows the use of high frame rates,
- imaging with polarization separating the left and right partial images L and R. They are displayed simultaneously. Lines of partial images L, R are interleaved (usually in the vertical direction). In the front of the display is placed a polarizing filter (Film-type Patterned Retarder). The viewer observes test images over the uncontrolled passive polarized glasses. A typical example is the LG CINEMA 3D system,
- auto-stereoscopic display which does not require any optical instruments (glasses). Its optical part arranges separation of light flow emitted from vertical strips of the partial images L and R, whereon the image is divided so that these partial flows strike only on the corresponding eye. It is realized in the form of a vertically oriented parallax grid, or more frequently as a set of vertical strips of lenticular lenses deposited in the front of the display.

5.2.1 TV sets selected for testing

For testing, the following reputable 3D TVs were used 3D TVs:

- 3D TV Panasonic TX-P42GTT20E with Full HD plasma display and active controlled LCD glasses,
- 42LW570S LG 3D TV with Full HD LCD display and passive polarized,
- 3D monitor Toshiba Qosmio F-750 with 15 " LCD 3D auto-stereoscopic display.

The auto-stereoscopic display is originally designed for a PC and therefore for one observer. It is equipped with a system for monitoring the position of the head DTH (Dynamic Head Tracking), in which the camera follows the position of the viewer's head and from this it optimizes the position of the 3D image on the display.

5.2.2 Measuring workplace

Objective measurements of photometric parameters and subjective testing of spatial perception and quality of 3D images reproduced by different displays were realized in different observer positions - for different visual angles. Arrangement of the workplace is in Fig. 5.1. Visual angle (α) is changed by 10°. Observer positions are placed on a circle. The optimum viewing distance (L = 1.8 m for TVs with 40" diagonal and 0.6 m for the 15" monitor) is maintained in all test positions. Due to

the axially symmetric predictable rated parameters testing was performed only in one direction. Subjective tests and the previously measured objective photometric parameters were carried out in a partially darkened room in order to reduce the effect of external lighting.

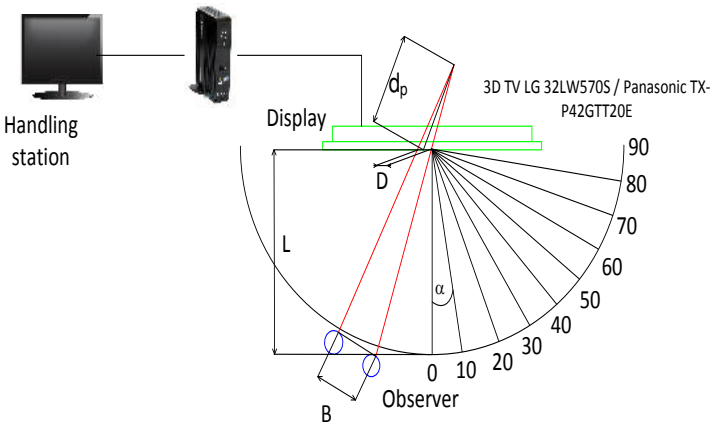


Fig. 5.1: Schematic arrangement of the workplace.

5.2.3 Measurement of photometric parameters of tested displays

Objective measurements of the directional dependence of photometric parameters of all three displays, which were used for 3D imaging, were also a component part of these tests. This, at least enables a partial evaluation of the impact of display technology on the subjectively assessed viewer's spatial perception. Electronic signals of four test patterns (red, blue, green and white area with 100% saturation), generated by the TV generator, were used for measuring. Displays were set to maximum saturation S and approximately the same brightness (for perpendicular direction). Measurements of photometric and colorimetric parameters were made with the Chroma Meter CS-100A Konica Minolta. The brightness B_i and saturation S_i of the three basic colors R, G, B for different angles α were determined from the measured trichromatic coordinates x_i and y_i using the CIE diagram. Subsequently, they were calculated relative to the values $S_{i,r} = Si(\alpha)/S_{i,0}$, relative to the maximum

value $S_{i,0}$, for perpendicular direction of measurement ($\alpha = 0$). The same method was used for measuring angles calculated by the relative brightness $B_r = B(\alpha)/B_0$. Measurement results are depicted in graphical form in Fig. 5.2, Fig. 5.3 and Fig. 5.4. They confirm the known fact that the directional dependence of color reproduction of plasma displays is less than LCD displays. However, it is apparent, that for the small visual angles α in the horizontal direction (up to the 20%) degradation of color reproduction for all three display types is approximately similar and relatively small.

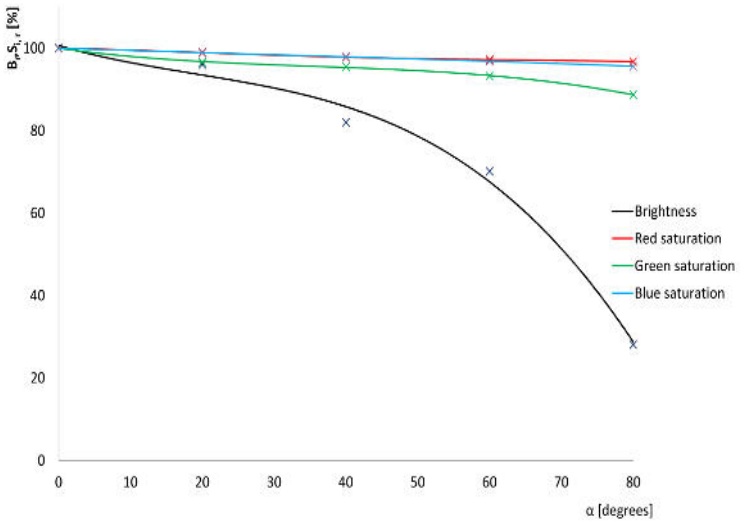


Fig. 5.2: Dependence of the relative color saturation S and brightness B on the viewing angle α for the plasma TV set Panasonic TX- P42GTT20E.

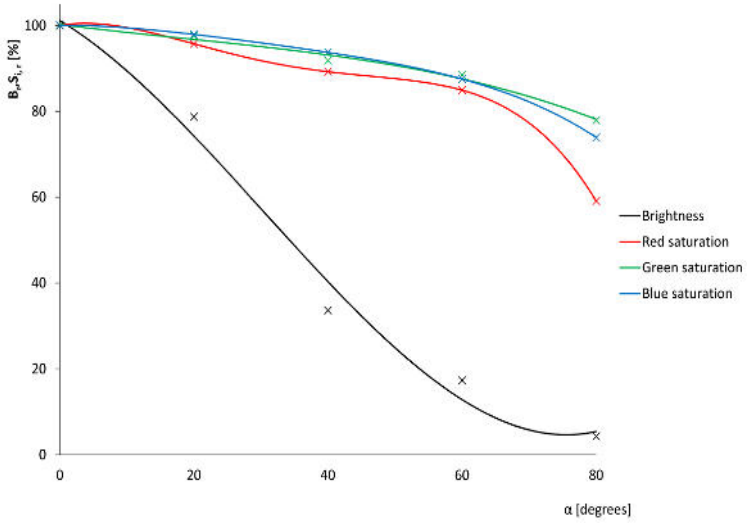


Fig. 5.3: Dependence of the relative color saturation S and brightness B on the viewing angle α for LCD TV set LG 42LW570S.

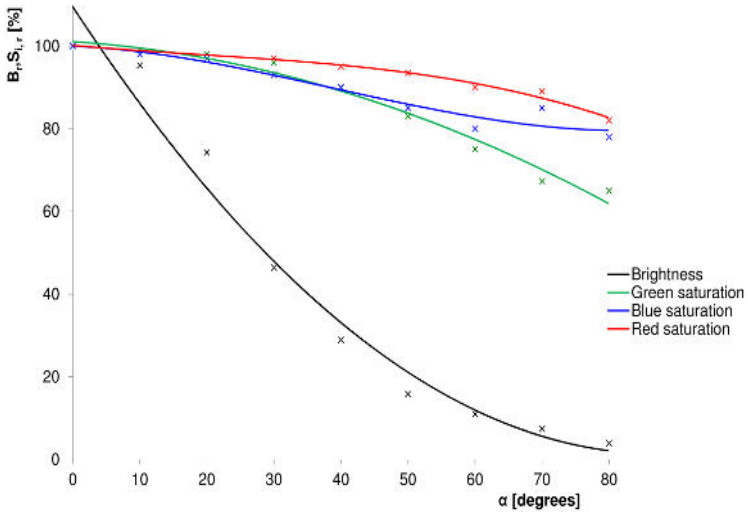


Fig. 5.4: Dependence of the relative color saturation S and brightness B on the viewing angle α for LCD 3D auto-stereoscopic 15" monitor Toshiba Qosmio F-750.

5.2.4 Testing methods

An observer evaluates the spatial perception (depth observed scene) in each position and subjectively perceived the quality of displayed 3D images - especially sharpness, 3D crosstalk, color rendering, motion distortion etc. The evaluator's visual abilities were verified by special check tests. For testing, we selected a group of 28 observers within an age range of 15 -70 years. Two different methods for evaluating perception were used.

- **Method A (without reference)** - observers use a seven-stage scale - (7- the best, 1 the worst) for immediate rating. Evaluators used a short questionnaire containing some sub-questions.
- **Method B (with references)** – evaluation in the previous position represents the reference for evaluating in the following position. Evaluation begins in the first position, which corresponds to the optimal (perpendicular) visual angle = 0). The evaluator in the following position evaluates both parameters by the percentage expression of the comparison of the image evaluation in the previous position.

Test results realized by method A were also converted to a percentage scale (7 matches 100%) for a uniform interpretation of results. Because the test results obtained by these two methods differ a little for statistical processing and final graphic display, their average is used. The results obtained by these various ways of evaluations were equal. This fact is one of the important findings of this experiment.

5.2.5 Used testing images and movies

The results of subjective tests may also be influenced by the content of the evaluated 3D images and video-sequences. Three 3D video sequences with a duration of about 15 seconds, containing a scene with different depth, were selected for testing. Three 3D still images selected from these video-sequences were also evaluated. Evaluators had approximately 5 seconds for each evaluation. All 3D image tests were obtained mainly from Blu Ray discs and played by the media player X-streamer Ultra in native Full HD resolution (1080 x1920 pixels) and in the "Side by Side" format. For test purposes, we intentionally did not use special 3D video sequences with unnatural depth of scenes that are scanned by 3D camera systems with variables and enlarged stereo-bases, which is also reflected in negative parallax (eg. Avatar, various computer games, etc.).

5.2.6 Results of the subjective tests

Subjective test results of directional dependence of 3D image quality and spatial perception are shown in graphical form in Fig. 5.5 and Fig. 5.6. The following is the color representation used for these figures: red: Panasonic with active glasses, green LG with passive glasses, blue: Toshiba auto-stereoscopic display.

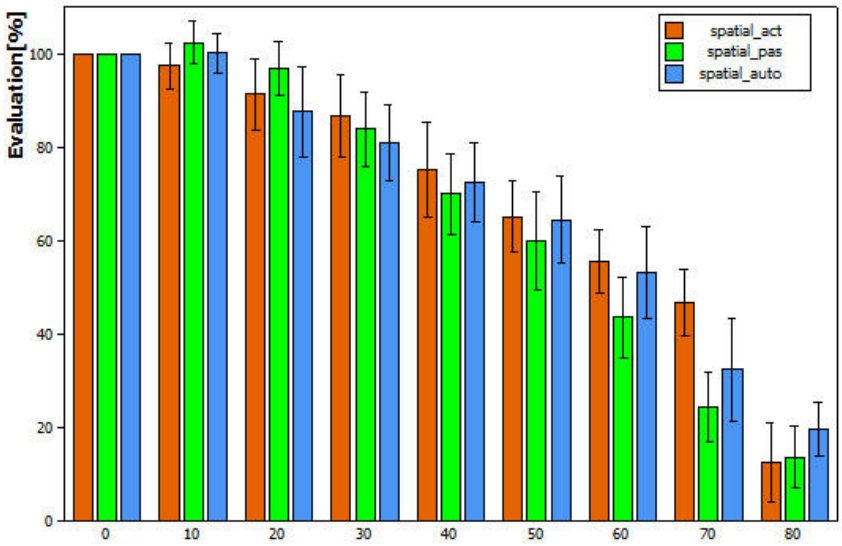


Fig. 5.5: Results of the subjective tests of the spatial perception dependency on view angle for 3D images.

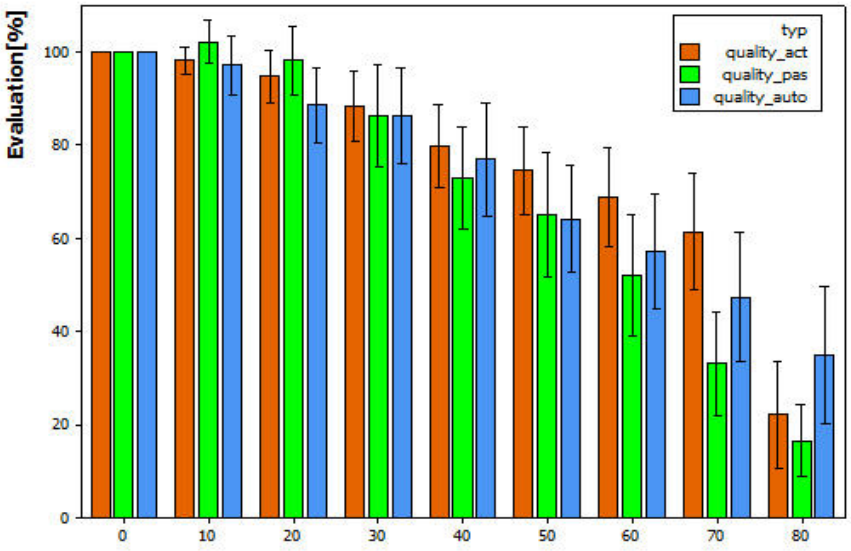


Fig. 5.6: Results of the subjective tests of the image quality dependency on view angle for 3D images.

5.2.7 Statistical processing of the subjective tests results

Subsequently, after performing the subjective tests, we executed a statistical analysis of the obtained results. For this purpose, we used the developed environment MATLAB and statistical software Minitab. One of the most important tasks was detecting outliers (odd results)[125], [126]. In this work, outliers are respondents whose evaluation distinctly deviates from the mean value in most viewing angles. The term outlier does not indicate evaluation for individual positions (viewing angle), but the respondent as a whole (evaluation in all viewing angles). The conditions for indicating a respondent as an outlier is deviation of their evaluation in most viewing angles. Therefore, the outliers were detected in the first step. In the beginning, we confirmed by using a test that the obtained subjective evaluation has Gaussian distribution. Consequently, we can utilize Grubbs test for detecting outliers. The Grubbs test was proposed by F.E. Grubbs [127]. The test is performed by the following equation

$$G = \frac{|x - \overline{x}|}{s}, \tag{5.1}$$

where x is the actual tested data point (in this case x is the numerically expressed evaluation of a certain respondent), s is the standard deviation of the data set and is the mean of the data set (in this case, the set of data are evaluations of all respondents at the relevant angle).

Subsequently, we compare result G with tabulated value for a given number of points in the dataset and demanded confidence (commonly a confidence of 95 is

used). If the calculated value G is bigger than the critical value for a given number of attempts (2.557 for 20 attempts), then the response is rated as an outlier. We find outliers separately in each of the following evaluated questions:

- spatial perception for an active display system, hereinafter referred to as spatial_act,
- spatial perception for passive display system,hereinafter referred to as spatial_pas,
- spatial perception for auto-stereoscopic display system,hereinafter referred to as spatial_auto,
- quality perception for active display system,hereinafter referred to as quality_act,
- quality perception for passive display system,hereinafter referred to as as quality_pas,
- quality perception for auto-stereoscopic display system,hereinafter referred to as quality_auto.

Detection of outliers is an important task. Therefore, it is advisable to test various methods for reliable detection. In the next step, we used cluster analysis. Evaluations in all viewing angles of each respondent separately for the above-mentioned questions present inputs for the following operation. Therefore, every respondent is described by nine variables (nine viewing angles). It is necessary to asses which respondent is distinctly various. Dendrogram is used for this purpose. Dendrogram is a convenient way of depicting pair-wise dissimilarity between objects, often used with the topic of cluster analysis. In other words, dendrogram expresses correlation of data. On the horizontal axis, individual respondents are arranged and on the vertical axis there is similarity for which an individual subject can be assigned to an appropriate cluster. Consequently, we can see and assess which respondent fits least to the set. For example, the dendrogram for spatial_act is shown in Fig. 5.7. We can see that the most dissimilar respondent is number C136 (C136 is only a working label) in this example. In the same way, we assess all evaluated questions. Besides the Grubbs test and cluster analysis, we also used the Principal Component Analysis (PCA) [128]. The main aim of PCA is to condense information, which is contained in a great number of original variables. Consequently, the result of PCA is a smaller number of variables with a minimal loss of information. PCA works with linear dependency of original variables and it defines new independent variables based on this dependency. Using PCA, we calculate new components (variables) which serve to the description of the results of the test. From the results of PCA, we can obtain further information which can help in detecting outliers [129], [130]. The detection of an outlier is executed by analyzing a biplot. A biplot allows you to visualize the magnitude and sign of each variable contribution to the first two

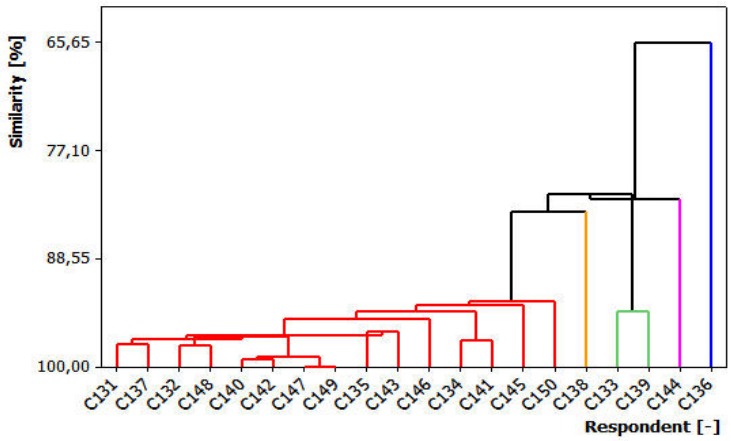


Fig. 5.7: Example of the dendrogram. Detection of the outliers in evaluating spatial effect on the active system spatial_act using dendrogram.

principal components, and how each observation is represented in terms of those components. Each observation (respondent) is represented by a descriptor. The angle between descriptors is proportional to the correlation between observations. The scree plot for spatial_act is shown in Fig. 5.8. Moreover, we can determine the number of required components for describing the data set by using a scree graph. A scree graph contains eigenvalues of each component. The point in which the curve begins to straighten represents the maximum number of components necessary for description. For example, a scree graph for spatial_act is shown in Fig. 5.8. We can sufficiently describe a set of evaluations from individual respondents by three components. This fact implies that the evaluation of respondents is mutually very similar. Great similarity is a mark that the results of subjective tests have good reliability. This is another proof that evaluations of individual respondent have some legality.

The results indicate that eliminating outliers does not distinctly improve the confidence interval of results in spatial effect. For evaluating the quality better, improvement of the confidence interval (about 2 percent for each angle) is reached. This fact is proof that evaluating spatial perception in 3D TV is individual in certain scales and we can not obtain results with a minimal confidence interval. Hence, there will always be some variance of evaluation of individual correspondents. However, a set answer can be described only by two or three PCA components. This means, that the evaluation of the individual respondents is relatively similar. Often, situations occur that the evaluation of some respondent differs significantly in one or two viewing angles. Consequently, this fact increases the confidence interval. Despite this, results of the test are reliable. We confirmed this fact by using the analysis ANOVA (Analysis of Variance)[131]. ANOVA is a statistical method for comparing

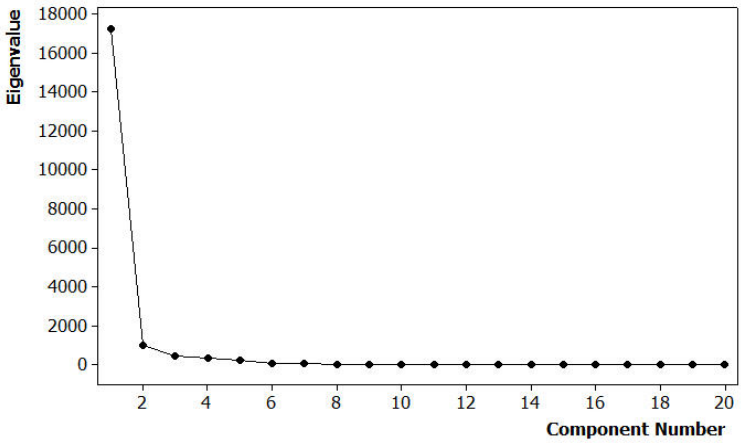


Fig. 5.8: Example of the PCA scree graph. PCA analysis of the spatial effect for the active system spatial_act.

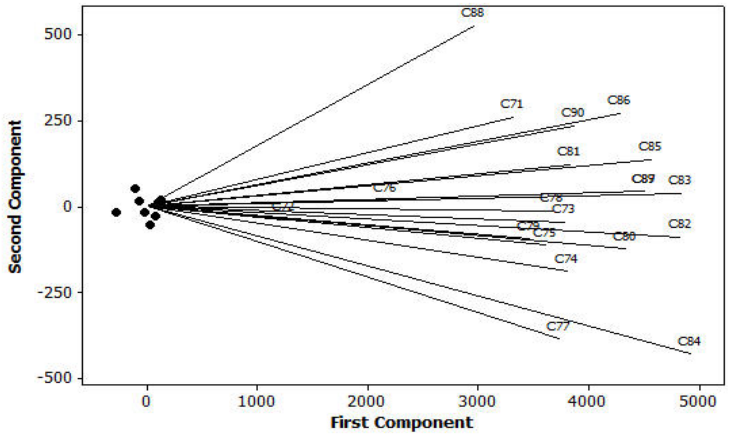


Fig. 5.9: Example of the PCA biplot. Detection of the outliers in evaluating the special effect on the active system.

the similarity of two or more sets of data. The main principle is assessment of the truthfulness of the null hypothesis. The null hypothesis says that every sample in the set is similar. In other words, between various samples there are no significant differences. In our case, the evaluations of individual respondents are similar and therefore the results of the test are reliable. The result of ANOVA is p-value. If p-value is near zero, then we can say that the null hypothesis is not right and that at least one sample mean is significantly different than other sample means. A common significance level is 0.05 or 0.01. ANOVA proved similarity of the respondents' answer to questions about spatial perception. On the contrary, the result of ANOVA refutes similarity of the respondents' answer to questions about quality. This situation can be caused by inaccurate definition of quality to respondents.

Evaluation	p-value	Truthfulness of null hypothesis (significance level is 0.05)
spatial_act	0.4939	Yes
spatial_pas	0.8603	Yes
spatial_auto	0.7291	Yes
quality_act	0.0954	Yes
quality_pas	0.0248	No
quality_auto	0.0122	No

Tab. 5.1: Results of ANOVA analysis with determining truthfulness of the null hypothesis.

viewing angle	0	10	20	30	40	50	60	70	80
spatial_act	0	4.98	8.56	7.52	8.59	8.73	9.48	10.89	6.31
spatial_pas	0	5.13	6.51	6.84	8.19	9.17	8.47	8.31	7.19
spatial_auto	0	2.25	3.86	6.55	9.67	7.08	6.42	7.71	9.41
quality_act	0	5.64	8.39	9.90	11.79	7.56	9.61	11.60	11.58
quality_pas	0	4.30	5.26	7.99	9.98	9.58	12.93	11.07	7.34
quality_auto	0	3.14	3.02	6.98	8.86	8.01	10.09	12.43	8.38

Tab. 5.2: Confidence intervals for all tested display and viewing angles.

5.2.8 Conclusion

This chapter deals with one aspect that affects the viewer's spatial perception when viewing 3D images reproduced on various types of 3D displays. The influence of the viewing angle on the resulting spatial perception and quality of the 3D images and video were evaluated using objective measurements and subjective tests. Tests were performed on three TV sets with different displays and methods of 3D displaying. The purpose of the objective measurements was to determine the dependency of the relative color saturation S and brightness B on the viewing angles for all three TV systems used and assess the relationship between this dependency and subjective evaluations. All tests were realized in the workplace displayed in Fig.5.1. The objective measurement confirmed the known fact, that the plasma display has a wider viewing angle. The dependency of the saturation and brightness on the viewing angle have slower progress for the plasma TV than for the LCD TV (see Fig.5.2, Fig.5.3 and Fig.5.4). Subjective tests were organized after the objective measure-

ments. Thirty-eight respondents of different ages evaluated two various parameters of the Quality Of Experience: spatial effect and image quality. We used two different ways of evaluation which are described in section 5.2.4. It was shown that different testing methods have negligible effect on the results of the evaluation tests and the results of both methods were therefore averaged. An important part of these tests represents the statistical assessment of the respondent's answer. In the first step, it was necessary to detect outliers. For this purpose, we used various methods: the Gruber test, PCA analysis and cluster analysis. Ways and conditions to detect outliers are described in more detail in section 5.2.7. The aim of the detection of outliers is to increase the reliability of the results. The confidence intervals for evaluation in each viewing angle were calculated before outlier detection and after outlier detection for every display system and all evaluated quality parameters. The results of the statistical analysis are summarized in Tabs. 5.2.7 and 5.2.7. Unfortunately, the values of the confidence interval are about 10 percent, especially at high viewing angles. Subsequently, it can be seen that in these tests, detection of outliers did not have significant impact on the confidence interval. The results of the tests also confirmed the assumption that evaluating spatial effect is affected by the observer's physiological and psychological features too. This may cause various variances in evaluation. The examination of this fact would be the object of further tests. Although we need to execute further tests, we can observe some important facts from the results. The evaluation of both parameters (spatial effect and image quality) is almost identical for small viewing angles. Subsequently, the evaluation begins to differ for higher angles and an active system is evaluated better than a passive system. This fact is in accordance with the results of the objective measurement.

6 CONCLUSION

In the introduction of my dissertation, there are four various aims which are logically related. A proposal of a new method for finding corresponding points was the first aim. The test compared commonly used methods (see section 2.2.4) preceded before my proposal of a new approach. The new method for detecting a corresponding point for a specific selected point was proposed in section 2.3. The proposed method is based on the model of probability of the movement of the points in the examined area of the image. The new approach can be used with an advantage especially for finding a corresponding point for points selected by the user in an area with small contrast. The proposed algorithm reaches a result with much better reliability than methods commonly used for this purpose. This fact is confirmed by the results of the experiment described in section 2.3.3. The main principle of this method was published in paper [136]. Extended and more detailed algorithms were published in paper [141]. The proposal of the method for finding corresponding points by conversion to pseudo colors is an important innovation described in section 2.4. The executed tests confirmed the usability of this method especially with an image area without contrast where finding points in true colors fails. The disadvantage of this method is a decrease in reliability. This problem can be solved by combining the monochromatic method and subsequent elimination of false correspondences. This approach was published in paper [136]. Both proposed methods were implemented to the designed software for reconstructing the model of the scene and calculating the depth maps (see section 2.5).

The next aim was to analyze the achievable accuracy of the model scene reconstruction. The analysis is logically connected with the previous examination dealing with finding corresponding points. Finding corresponding points is a fundamental step in the process of reconstructing a spatial model of the scene, which critically influences the accuracy of the reconstruction. Two views on accuracy were used in this work. The first aim was to extend previous work in the area of investigation of the effects of camera alignment errors on estimating depth by a stereoscopic camera system. The evaluation of effect of the investigated phenomena on the remaining spatial coordinates extend previous work in the area of investigation of the effects of camera alignment errors on estimating depth by a stereoscopic camera system. The effect of the errors in camera alignment on the remaining spatial coordinates was investigated. The practical experiment revealed that the previously derived equation for error in depth due to errors in camera rotation are incorrect. Therefore, a new equation for error in depth estimation was derived and its correctness was proven by experiment. Subsequently, the equations for estimation error in the remaining two spatial coordinates were derived and their correctness was proven.

The relative errors are used for presenting the results. The relative expression has greater informative value then the absolute. The second part deals with the impact of inaccurate determination of corresponding points. This impact is investigated in both camera systems: stereo alignment and even universal arrangement. The error is expressed directly for a stereo system. The investigation of the universal arrangement connects the error in corresponding points with the error in camera alignment. Exterior calibration is considered as the transformation between these two camera systems. Then, the consequence of error in the rotation matrix is considered as the error in the resulting stereo alignment. The error in the rotation matrix is caused by inaccurate determination of corresponding points. The impact of various errors in finding corresponding points on the error of the exterior calibration is a very complex problem with a large number of degrees of freedom. Therefore, the statistical probabilistic analysis is the sole solution. However, this analysis is useful and important, because it shows how big an error can occur. The results, among others, confirm that corresponding points is the most important step of the reconstruction. A small error in this step caused a critical error in reconstruction. My two papers on the topic of accuracy are in review process.

The creation of the depth maps closely relates to the reconstruction of the model of the scene, especially with finding corresponding points. The depth map represents one of the important forms of describing the 3D scene. The depth maps are suitable, for example, for transmitting the 3D signal. If the inputs for their estimation are two 2D images shifted only about stereo base, then the depth map is theoretically given by the horizontal parallax of all the individual points. I proposed three methods for generating depth maps. Two of them are based on using stereo images (see section 4.1), therefore they belong to the group passive methods. It is based on generating an initial depth map by using commonly used methods, for example SAD. Subsequently, the initial depth map is improved by two various methods. The first of them is based on edge representation and using spatial continuity of the depth map. The second one is based on image segmentation and using correspondences of the significant points found by the algorithm SURF. These methods were published in paper [139]. The third approach is a system based on using the combination of a passive stereoscopic method and an active optical incoherent active scanning method (see section 4.2). The proposed approach uses the advantage of both methods. The two maps are generated and subsequently transformed to one. This method was published in paper [135].

The last but not least important aim of my work was executing and evaluating subjective tests of the spatial effect of the stereoscopic 3D videosequence on the technologically different 3D displays with different 3D systems. The subjective spatial perception and quality of the 3D image is influenced by many objective and sub-

jective aspects: the content of the scene and quality of the image (video sequence), system of the 3D imaging, technology of the display, viewing condition (e.g. room illumination, viewing angle and distance) physiological and psychological status of the viewer. These aspects cannot be separated during the test. It is necessary to constantly hold every aspect except one which we examined. In the first test, the influence of changing the viewing angle in both directions was examined. The tests were executed separately on various displays (CCD and plasma) with same diagonal and same native resolution (FULL HD) with various 3D systems. The measurement of the objective parameters of the displays was executed before the subjective tests. For executing the tests, a methodology for testing was proposed. I executed statistical analyses of the results (see section 5.2.7). The results were published in paper [140]. Other aspects which have an influence on spatial perception were examined in an experiment which was executed by a large research team which I was a part of.

BIBLIOGRAPHY

[1] ZHANG, Z. Flexible camera calibration by viewing a plane from unknown orientations. *Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on* , vol.1, no., pp.666-673 vol.1, 1999 doi: 10.1109/ICCV.1999.791289.

[2] FAUGERAS, O.D., LUONG, Q.T., AND MAYBANK, S. Camera self-calibration: theory and experiments. *In Proc. European Conference on Computer Vision, LNCS588*, pages 321–334. Springer-Verlag, 1992.

[3] FAUREGAS, O. What can be seen in three dimensions with an uncalibrated stereo rig? I*n Proccedings of European Conference on Computer Vision*, pages 563-578. Springer-Verlag, 1992.

[4] HARTLEY, R., GUPTA, R. AND CHANG, T. Stereo from uncalibrated cameras. *In Proceedings of international Conference on Computer Vision and Pattern Recognition*, Urbana Champaign, IL, USA, IEEE Comput. Soc. Press, pages 761-764, 1992.

[5] H LUONG, Q.-T., FAUREGAS, O. The fundamental matrix: theory, algorithms, and stability analysis. *International Journal of Computer vision*, volume17, pages 589-599, 1994.

[6] ZELLER, C., FAUREGAS, O. Camera self-calibration, form video sequences: the Kruppa equation revisited. *Research report 2793*, INRIA, France, 1996.

[7] PONCE, J., MARIMONT, D., CASS, T. Analytical methods for uncalibrated stereo and motion reconstruction. I*n Proccedings of European Conference on Computer Vision*, pages 463-470. 1994.

[8] BOUFAMA, B., MOHR, R. Epipole and fundamential matrix estimation using the virtual parallax property. *In Proccedings of IEEE International Conference on Computer Vision*, pages 1030-1036, Boston, MA, 1995.

[9] CHAI, J., MA, S. Robust epipolar geometry estimation using generic algorithm. *Patter Recognition Letters*, 19(9):829-838, 1998.

[10] TORR, P.H.S., MURRAY, D.W. The development and comparison of robust methods for estimating the fundamental matrix. *Int. Journal of Computer vision*, vol. 24, no.3, pp.271-300, 1997.

[11] ZHANG, Z., DERICE, R, FAUREGAS, O., LUONG, Q.-T. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. *Artificial Intelligence*, vol. 78,pp. 87-119, 1995.

[12] QUAN, L. Affine stereo calibration for relative affine shape reconstruction. *In Proceeding of British Machine Vision Conference*, pp. 659-668, 1993.

[13] FAUREGAS, O. Stratification of three-dimensional vision: projective, affine, and metric representation. *Journal of the Optical Society of America*, vol. 12, pp. 465-484, 1995.

[14] STURM, P. Critical motion sequences for monocular self-calibration and uncalibrated Euclidean reconstruction. *In Proceedings of international Conference on Computer Vision and Pattern Recognition*, pp. 1000-1005, 1997.

[15] STURM, P. Critical motion sequences for self-calibration of cameras and stereo systems with variable focal length. *In Proceeding of British Machine Vision Conference*, pp. 63-72, 1999.

[16] TORR, P.H.S., FITZGIBBON, A., ZISSERMAN, A. The problem of degeneracy in structure and motion recovery from uncalibrated images sequences. *Int. Journal of Computer vision*, volume 32,pp. 27-44, 1999.

[17] COLLINS, R., WEISS, R. Vanishing point calculation as a statistical inference on the unit sphere. I*n Proccedings of IEEE International Conference on Computer Vision*, pp. 400- 403, 1990.

[18] LUTTON, E., MAITRE, H., LOPEZ-KRAHE, J. Contribution to the determination of vanishing points using Hough transformation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no., pp. 430-438, 1994.

[19] MARR, D. AND POGGIO, T. A computational theory of human stereo vision. *Proc. R. Soc.*, pp. 263-295, 1979, 978-1-4684-6777-2.

[20] POLLARD, S.B., MAYHEW, J.E.W, FRISBY, J.P. PMF: A stereo correspondence algorithm using a disparity gradient constraint. *Perception*, vol.14, pp.449-470, 1985.

[21] BAKER, H.H., BINFORD, T.O. Depth from edge- and intensity-based stereo. *In Proceedings 7th Joint conference on Artificial Intelligence, Vancouver, Canada*, pp. 631-636, August 1981.

[22] BELHUMEUR, P. N. A bayesian approach to binocular stereopsis. *International Journal of Computer Vision (IJCV)*, vol.19, issue.3, pp. 237–262, 1996, ISSN: 0920-5691.

[23] OHTA, Y., KANADE, T. Stereo by intra- and inter-scanline search using dynamic programming. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 7, issue. 2, pp. 139– 154, 1985, doi: 10.1109/TPAMI.1985.4767639.

[24] SZELISKI, R. *Computer Vision algorithm and applications*. Springer 2011,812 pages, ISBN 978-1-84882-934-3.

[25] HAMZA, R.A., RAHIM, R.A., NOH, Z.M. Sum of Absolute Differences algorithm in stereo correspondence problem for stereo matching in computer vision application. *Computer Science and Information Technology (ICCSIT), 2010 3rd IEEE International Conference on* , vol.1, no., pp.652,657, 9-11 July 2010,doi: 10.1109/ICCSIT.2010.5565062.

[26] SCHMID, C., MOHR, R., BAUCKHAGE, C. Evaluation of interest point detectors, *International Journal of Computer Vision*, vol. 37, no.2, pp. 151–172, 2000,ISSN 0920-5691.

[27] RODEHORST, V., KOSCHAN, A. Comparison and Evaluation of Feature Point Detectors. *In Proc of the 5th Int Symposium TurkishGerman Joint Geodetic Days TGJGD 2006*, L. Gründig and M. O. Altan, eds. (Citeseer), p. 1-8.

[28] GALES, G., CROUZIL, A., CHAMBON, S. Complementarity of feature point detectors. *In International Conference on Computer Vision Theory and Applications VISAPP 2010*,Angers, France, May 17-21, 2010.

[29] BAY, H., TUYTELAARS, T., GOLL. L.V. SURF: Speeded up robust features. *Computer Vision and Image Understanding*, vol. 110, issue 3, June 2008, pp. 346-359, ISSN 1077-3142.

[30] LOWE. D. Distinctive Image Features from Scale-Invariant Keypoints. *International Journal of Computer Vision*, vol. 60, issue 2, pp.91-110, 2004.

[31] HARRIS, C., STEPHENS, M. A. combined corner and edge detector. *Proc. 4 the Alvey Vision Conference*, pp. 147-151, 1988.

[32] PEIPONEN, K.-E., Myllyla, R.,Priezzhev, A.V. *Optical Measurement Techniques: Innovations for Industry and the Life Sciences*. Springer, 2009. 155p., ISBN 9783540719267.

[33] HARDING, K. *Handbook of Optical Dimensional Metrology*. CRC Press, 2013. p. 492, ISBN: 9781439854815.

[34] NORGIA, M., GIULIANI, G., DONATI, S. New absolute distance measurement technique by self-mixing interferometry in closed loop. *Instrumentation and Measurement Technology Conference, 2004. IMTC 04. Proceedings of the 21st IEEE* , vol.1, no., pp.216,221 Vol.1, 18-20 May 2004 doi: 10.1109/IMTC.2004.1351031

[35] TAKASAKI, H. Moiré topography. *Applied Optics*, vol. 9, issue 6, pp. 1467-1472, 1970.

[36] CHIANG, FP. Moiré method for contouring displacement, deflection, slope, and curvature. *Proceedings of SPIE 153. Advances in optical metrology*, vol. II: 1978: 113–9.

[37] IDESAWA, M., YATAGAI, T., SOMA, T. Scanning moiré method and automatic measurement of 3-D shapes. *Applied Optics*, vol. 16, issue 8, pp. 2153–2162,1970.

[38] SU, X. Fourier transform profilometry: a review. *Optics and Lasers in Engineering*. Vol. 35, issue 5, pp. 263–284, May 2001.

[39] JINGANG, Z., JIAWEN, W. Spatial Carrier-Fringe Pattern Analysis by Means of Wavelet Transform: Wavelet Transform Profilometry. *Applied Optics*. Vol. 43, pp. 4993-4998, 2004.

[40] PATIL, A., RASTOGI, P. Approaches in generalized phase shifting interferometry. *Optics and Lasers in Engineering*. 2005, Vol. 43, pp. 475-490. 2005.

[41] QINGYING, H., HARDING, K.G. Conversion from phase map to coordinate: Comparison among spatial carrier, Fourier transform, and phase shifting methods. *Optics and Lasers in Engineering*, Vol. 45, issue 2, pp. 342-348, February 2007, ISSN 0143-8166.

[42] AYACHE, N., HANSEN, C. Rectification of images for binocular and trinocular stereovision. *Pattern Recognition 9th International Conference on*. Vol. 1, pp. 14-17, Nov 1988.

[43] LIEBOWITZ, D., ZISSERMAN, A. Metric rectification for perspective images of planes. *Computer Vision and Pattern Recognition Proceedings IEEE Computer Society Conference on*. Vol., no., pp.482-488, 23-25 Jun 1998.

[44] SCHARSTEIN, D. SZELISKI, R. A taxonomy and evaluation of dense twoframe stereo correspondence algorithms. *Technical Report MSR-TR-2001-81, Microsoft Corporation*. Redmond, WA 98052, USA, 2001.

[45] BROWN, M.Z., BURSCHKA, D., HAGER, G.D. Advances in computational stereo. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*. vol. 25, no.8, pp. 993- 1008, Aug. 2003 doi: 10.1109/TPAMI.2003.1217603.

[46] KYTÖ, M., NUUTINEN, M., OITTINEN, P. Method for measuring stereo camera depth accuracy based on stereoscopic vision. *Proceedings of SPIE/IS&T Electronic Imaging, Three-Dimensional Imaging, Interaction, and Measurement*. San Francisco, California, USA, 24.-27.1.2011. ISBN: 9780819484017. 9 p.

[47] HARRIS, J.M. Monocular zones in stereoscopic scenes: A useful source of information for human binocular vision? *Stereoscopic Displays and Applications XXI*, vol. 7524, pp. 11. 2010.

[48] TAO, Z., BOULT, T. Realistic stereo error models and finite optimal stereo baselines. *Applications of Computer Vision (WACV), 2011 IEEE Workshop on.* pp.426-433, 5-7 Jan. 2011 doi: 10.1109/WACV.2011.5711535.

[49] ZHAO, W., NANDHAKUMAR, N. Effects of camera alignment errors on stereoscopic depth estimates, *Pattern Recognition.* Vol. 29, issue. 12, December 1996, pp. 2115-2126, ISSN 0031-3203, 10.1016/S0031-3203(96)00051-9.

[50] CHANG, C.; CHATTERJEE, S. Quantization error analysis in stereo vision. *Signals, Systems and Computers, Conference Record of The Twenty-Sixth Asilomar Conference on*. Vol.2, pp.1037-1041, 26-28 Oct 1992. doi: 10.1109/ACSSC.1992.269140

[51] STANČIK, P. *Optoelektronické a fotogrammetrické měřící systémy*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2008. 89p. Supervisor of of the dissertation prof. Ing. Václav Říčný, CSc.

[52] GALLUP, D., FRAHM, J.-M., MORDOHAI, P., POLLEFEYS, M. Variable baseline/resolution stereo. *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*. vol., no., pp.1-8, 23-28 June 2008. doi: 10.1109/CVPR.2008.4587671

[53] BELHAOUA, A., KOHLER, S., HIRSH, E. Error Evaluation in a Stereovision-Based 3D Reconstruction Systém. *Image Video Process*. Article 2, 12 pp, 2010.

[54] BELHAOUA, A., KOHLER, S., HIRSH, E. Estimation of 3d reconstruction errors in a stereo-vision system. *In Proceedings Modeling Aspects in Optical Metrology II, vol. 7390 of Proceedings of the SPIE.* pp. 1–10, Optical Metrology, Münich, Germany, June 2009.

[55] BELHAOUA, A., KOHLER, S., HIRSH, E. Estimation of 3d reconstruction errors in a stereo-vision system. *In Proceedings Modeling Aspects in Optical Metrology II, vol. 7390 of Proceedings of the SPIE*, pp. 1–10, Optical Metrology, Münich, Germany, June 2009.

[56] SWAN, J. E. II, LIVINGSTON, M.A., SMALLMAN, H.S., BROWN, D., BAILLOT, Y., GABBARD J.L., HIX, D. A. Perceptual Matching Technique for Depth Judgments in Optical, See-Through Augmented Reality. *In Proceedings of the IEEE conference on Virtual Reality (VR '06). IEEE Computer Society, Washington, DC, USA* 19-26. DOI=10.1109/VR.2006.13 http://dx.doi.org/10.1109/VR.2006.13.

[57] LE CALLET, P., MÖLLER, S., PERKIS, A. Qualinet White Paper on Definitions of Quality of Experience . European Network on Quality of Experience in Multimedia Systems and Services (COST Action IC 1003),2012.

[58] CHEN, W., FOURNIER, J., BARKOWSKY, M., LE CALLET, P. New Requirements of subjective video quality assessment methodologies for 3DTV. *Proc. 5th Int. Workshop Video Process. Quality Metrics (VPQM)*, 2010.

[59] QUAN, H-T., LE CALLET, P., BARKOWSKY, M. Video quality assessment: From 2D to 3D — Challenges and future trends. *Image Processing (ICIP) 17th IEEE International Conference on*, pp.4025-4028, 2010,doi: 10.1109/ICIP.2010.5650571.

[60] LAMBOOIJ, M., IJSSELSTEIJN, W., BOUWHUIS, D.G., HEYNDERICKX, I. Evaluation of Stereoscopic Images: Beyond 2D Quality. *Broadcasting, IEEE Transactions on.* vol.57, no.2, pp.432-444, 2011, doi: 10.1109/TBC.2011.2134590.

[61] JOVELURU, P., MALEKMOHAMADI, H., FERNANDO, W.A.C., KONDOZ, A.M. Perceptual Video Quality Metric for 3D video quality assessment. *3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON).* pp.1-4,2010, doi: 10.1109/3DTV.2010.5506331.

[62] DE SILVA, V., FERNANDO, A., WORRALL, S., ARACHCHI, H.K., KONDOZ, A. Sensitivity Analysis of the Human Visual System for Depth Cues

in Stereoscopic 3-D Displays. *Multimedia, IEEE Transactions on.* vol.13, no.3, pp.498-506,2011, doi: 10.1109/TMM.2011.2129500.

[63] YASAKETHU, S.L.P., FERNANDO, W.A.C., KAMOLRAT, B., KONDOZ, A. Analyzing perceptual attributes of 3d video. *Consumer Electronics, IEEE Transactions on.* Vol.55, no.2, pp.864-872, 2009, doi: 10.1109/TCE.2009.5174467.

[64] *Subjective methods for assessment of stereoscopic 3DTV systems*, ITU-Recommendation BT.2021, 2012.

[65] *Subjective assessment of stereoscopic television pictures*, ITU-Recommendation BT.1438, 2000.

[66] *BT.2088 Stereoscopic Television*,Report ITU-R, 2006.

[67] KIM, D., MIN, D., JUHYUN, O., JEON, S., SOHN, K. Depth map quality metric for three-dimensional video. *Proc. Stereoscopic Displays and Applications XX.* Vol. 7237 ,2009, doi:10.1117/12.806898.

[68] SARIKAN, S.S., OLGUN, R.F., AKAR, G.B. Quality evaluation of stereoscopic videos using depth map segmentation. *Quality of Multimedia Experience (QoMEX), Third International Workshop on.* pp.67-71,2011,doi: 10.1109/QoMEX.2011.6065714.

[69] LIYUAN, X., JUNYONG, Y., EBRAHIMI, T., PERKIS, A. An objective metric for assessing quality of experience on stereoscopic images. *Multimedia Signal Processing (MMSP), IEEE International Workshop on.* pp.373-378,2010,doi: 10.1109/MMSP.2010.5662049.

[70] LIYUAN, X., JUNYONG, Y., EBRAHIMI, T., PERKIS, A. A perceptual quality metric for stereoscopic crosstalk perception. *Image Processing (ICIP), 17th IEEE International Conference on.* pp.4033-4036, 2010, doi: 10.1109/ICIP.2010.5649402.

[71] LIYUAN, X., JUNYONG, Y., EBRAHIMI, T., PERKIS, A. Estimating quality of experience on stereoscopic images. *Intelligent Signal Processing and Communication Systems (ISPACS), International Symposium on.* pp.1-4, 2010, doi: 10.1109/ISPACS.2010.5704599.

[72] HANHART, P., EBRAHIMI, T. Quality assessment of a stereo pair formed from decoded and synthesized views using objective metrics. *3DTV-Conference: The True Vision - Capture, Transmission and Display of 3D Video (3DTV-CON* pp.1,4, 2012, doi: 10.1109/3DTV.2012.6365478.

[73] HANHART, P., DE SIMONE, F., EBRAHIMI, T. Quality assessment of asymmetric stereo pair formed from decoded and synthesized views. *Quality of Multimedia Experience (QoMEX), Fourth International Workshop on.* pp.236-241, 2012, doi: 10.1109/QoMEX.2012.6263854.

[74] BOSC, E., PEPION, R., LE CALLET, P., KOPPEL, M., NDJIKI-NYA, P., PRESSIGOUT, M., MORIN, L. Towards a New Quality Metric for 3-D Synthesized View Assessment. *Selected Topics in Signal Processing, IEEE Journal of.* vol.5, no.7, pp.1332-1343, 2011, doi: 10.1109/JSTSP.2011.2166245.

[75] BOSC, E., KOPPEL, M., PEPION, R., PRESSIGOUT, M., MORIN, L., NDJIKI-NYA, P., LE CALLET, P. Can 3D synthesized views be reliably assessed through usual subjective and objective evaluation protocols? *Image Processing (ICIP), 18th IEEE International Conference on.* pp.2597-2600, 2011, doi: 10.1109/ICIP.2011.6116196.

[76] JIAN, Ch., HAIPENG, C., AUCHU, A.P., LAIDLAW, D.H. Effects of Stereo and Screen Size on the Legibility of Three-Dimensional Streamtube Visualization. *Visualization and Computer Graphics, IEEE Transactions on.* Vol.18, no.12, pp.2130-2139, 2012,doi: 10.1109/TVCG.2012.216.

[77] DE BOUGRENET DE LA TOCNAYE, J.L, COCHENER, B., FERRAGUT, S., IORGOVAN, D., FATTAKHOVA, Y., LAMARD, M. Supervised Stereo Visual Acuity Tests Implemented on 3D TV Monitors. *Display Technology, Journal of*, vol.8, no.8, pp.472-478, 2012, doi: 10.1109/JDT.2012.2198792.

[78] IJSSELSTEIJN, W. A, DE RDDER, H., VLIEGEN, J. Subjective Evaluation of Stereoscopic Images: Effects of Camera Parameters and Display Duration. *IEEE Transactions on circuits and systems for video technology.* vol. 10, no. 2, pp.225-233, 2000, DOI.10.1109/76.825722.

[79] YAMANOUE, H., OKUI, M., YUYAMA, I. A study on the relationship between shooting conditions and cardboard effect of stereoscopic images. *Circuits and Systems for Video Technology, IEEE Transactions on.* vol.10, no.3, pp.411-416, 2000, doi: 10.1109/76.836285.

[80] MIKHAIL, E.M., BETHEL J.S., McGLONE, J.CH. *Introduction to Modern Photogrammetry.* New York : John Wiley & Sons, 2001. 479 p. ISBN 0-471-30924-9.

[81] KRAUS, K. *Photogrammetry: Geometry from Images and Laser Scans.* 2nd edition. Berlín : Walter de Gruyter, 2007. 459 p. ISBN 978-3-11- 019007-6.

[82] MA, Y., SOATTO, S, KOSECKA, J., SASTRY, S.S. *An Invitation to 3-D Vision: From Images to Geometric Models*. 1st. Springer, 2003, 526 s, ISBN-10: 0387008934.

[83] Camera Calibration Toolbox for Matlab [open source software]: Jean-Yves BOUGUET, Last updated July 9th, 2010.

[84] HEIKKILA, J., SILVEN, O. A four-step camera calibration procedure with implicit image correction. *Computer Vision and Pattern Recognition, 1997. Proceedings., 1997 IEEE Computer Society Conference on*, pp.1106,1112, 17-19 Jun 1997, doi: 10.1109/CVPR.1997.609468.

[85] ABDEL-AZIZ, Y.I., KARARA, H.M. Direct linear transformation from comparator coordinates into object space coordinates in close-range photogrammetry. *Proceedings of the Symposium on Close-Range Photogrammetry. Falls Church, VA: American Society of Photogrammetry*,vol. 1.

[86] LONGUET-HIGGINS, H.C. A computer algorithm for reconstructing a scene from two projection. *Nature*, vol.293, pp 133-135, September 1981, doi:10.1038/293133a0.

[87] HARTLEY, R. I., STURM, P. Triangulation. *6th International Conference, CAIP'95, Prague, Czech Republic, September 6–8, 1995 Proceedings*, vol. 970, pp.190-197, 1995, ISBN 978-3-540-60268-2.

[88] VIOLA, P., JONES, M. Rapid Object Detection using a Boosted Cascade of Simple Features. *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, pp.511–518, 2001, doi: 10.1109/CVPR.2001.990517.

[89] HIRSCHMÜLLER, H., SCHARSTEIN, D. Evaluation of cost functions for stereo matching. *In IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2007)*,pp. 1-8 , June 2007, doi: 10.1109/CVPR.2007.383248.

[90] ZHOU WANG, BOVIK, A.C., SHEIKH, H.R., SIMONCELLI, E.P. Image quality assessment: from error visibility to structural similarity. *Image Processing, IEEE Transactions on*, vol.13, no.4, pp.600,612, April 2004,doi: 10.1109/TIP.2003.819861.

[91] NODA, I., OZAKI, Y. *Two-Dimensional Correlation Spectroscopy: Applications in Vibrational and Optical Spectroscopy*. John Wiley Sons, 2005, ISBN 9780470012390.

[92] CANNY, J. A Computational Approach To Edge Detection. *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 8, issue 6, pp. 679–698, 1986, doi: 10.1109/TPAMI.1986.4767851.

[93] GONZALES, R.C., WOODS, R.E., EDDINS, S.L. *Digital Image Processing Using MATLAB*. New Jersey, Prentice Hall, 2009, ISBN-13: 978-0-9820854-0-0.

[94] TSAI D-Y.,LEE, Y., MATSUYAMA, E. Information Entropy Measure for Evaluation of Image Quality. *Journal of Digital Imaging*, vol. 21, issue 3, pp 338-347,2008, DOI> 10.1007/s10278-007-9044-5.

[95] CHUAN, L., JINJIN, Z., CHUANGYIN, D., HONGJUN, Z. A. Method of 3D reconstruction from image sequence. *In 2nd International Congress on Image and Signal Processing (CISP)*, pp.1-5,2009., doi: 10.1109/CISP.2009.5305647.

[96] QUWEIDER, M.K., Adaptive Pseudocoloring of Medical Images Using Dynamic Optimal Partitioning and Space-Filling Curves. *Biomedical Engineering and Informatics, 2009. BMEI '09. 2nd International Conference on*, vol., no., pp.1-6, Oct. 2009, doi: 10.1109/BMEI.2009.5304855. doi: 10.1109/BMEI.2009.5304855.

[97] ZAHEDI, Z., SADRI, S., SOLTANI, M., TEHRANI, M.K. Breast diseases detection and pseudo-coloring presentation for gray infrared breast images. *Communications and Photonics Conference and Exhibition, 2011. ACP*, vol., no., pp.1-8, Nov. 2011, doi: 10.1117/12.905604.

[98] ABIDI, B.R., Yue, Z., GRIBOK, A.V. ABIDI, M.A. Improving Weapon Detection in Single Energy X-Ray Images Through Pseudocoloring. *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on*, vol.36, no.6, pp.784-796, Nov. 2006, doi: 10.1109/TSMCC.2005.855523. doi: 10.1109/TSMCC.2005.855523

[99] WANG, T., Su, J., HUANG, Y., ZHU, Y. Study of the pseudo-color processing for infrared forest-fire image. *Future Computer and Communication (ICFCC), 2010 2nd International Conference on* vol.1, no., pp.415-478, 21-24 May 2010, doi: 10.1109/ICFCC.2010.5497756.

[100] TWIDDY, R., CAVALLO, J., SHIRI, S.M. Restorer: a visualization technique for handling missing data. *Visualization 1994, Proceedings., IEEE Conference on* , vol., no., pp.212-216, 17-21 Oct 1994 doi: 10.1109/VISUAL.1994.346317.

[101] LEHMAN, T., KASER, A., REPGES, R. A simple parametric equation for pseudocoloring grey scale images keeping their original brightness progression. *Image and Vision Computing*. Vol. 15, issue 3, pp. 251–257, 1997, ISSN 0262-8856.

[102] LU, X., DING, M., WANG, Y. A New Pseudo-color Transform for Fibre Masses Inspection of Industrial Images. *Acta Automatica Sinica*. Vol. 35, issue 3,pp 233-238, 2009,ISSN 1874-1029.

[103] YOUVAN, D. Pseudocolor in Pure and Applied Mathematics : a Free on-Line e-Book with Source Code [online]. 2006., 1.1.2011 [cit. 2011-04-18]. <http://www.youvan.com/>. ISBN 978-0-615-43573-2.

[104] STRECHA, C., FRANSENS, R., Van GOOL L. Combined Depth and Outlier Estimation in Multi-View Stereo. *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, vol. 2, pp. 2394-2401, ISBN 0-7695-2597-0.

[105] CRAIG, J. *Introduction to Robotics: Mechanics and Control*. 3rd. Prentice Hall, 2004. 480 p. ISBN 0201543613.

[106] HERMANN, S., VAUDREY, T. The gradient - A powerful and robust cost function for stereo matching. *Image and Vision Computing New Zealand (IVCNZ), 2010 25th International Conference*, vol., no., pp.1,8, 8-9 Nov. 2010, doi: 10.1109/IVCNZ.2010.6148804.

[107] LANKTON, S. 3D Vision with Stereo Disparity. In: [online]. [cit. 2013-02-26]. Url: http://www.shawnlankton.com/2007/12/3d-vision-with-stereo-disparity/

[108] OSTEN, W., REINGARD, N. *Optical imaging and Metrology: Advanced technologies*, John Wiley Sons, 2008.

[109] KREIS, T. *Handbook of holographic Interferometry : Optical a digital methods.* Weinheim : WILEY-VCH Verlag GmbH Co. KGaA, 2005. 535 s. ISBN 3-527-40546-1.

[110] SCHARSTEIN, D. *View Synthesis using Stereo Vision*. Ph.D Thesis, Cornell University,1998.

[111] http://stereotracer.en.softonic.com/.

[112] HU, Q., HARDING, K. G. Conversion from phase map to coordinate: Comparison among spatial carrier, Fourier transform, and phase shifting methods

map to coordinate. *Optics and Lasers in Engineering* Vol. 45, issue 2, pp. 342–348, February 2007.

[113] XIANYU, S., WENJING, Wenjing, CH. Fourier transform profilometry: : a review. *Optics and Lasers in Engineering*. Vol. 35, no. 5, 2001,pp. 263-284, ISSN 01438166.

[114] REDMAN, B. Stand-off Biometric Identification using Fourier Transform Profilometry for 2D+3D Face Imaging. *Applications of Lasers for Sensing and Free Space Communications, OSA Technical Digest (CD) (Optical Society of America, 2011*, paper LThB3., pp. 3-5, 2011.

[115] HUI, T.-W., PANG G., 3D profile reconstruction of solder paste based on phase shift profilometry. *Industrial Informatics, 2007 5th IEEE International Conference on*, vol.1, no., pp. 165-170, 2007, doi: 10.1109/INDIN.2007.4384750.

[116] YEN, H.-N., TSAI, D.-M., YANG J.-Y. Full-field 3D measurement of solder pastes using LCD-based phase shifting techniques. *IEEE Transactions on Electronics Packaging Manufacturing*. vol. 29, no. 1, pp. 50-57, 2006, doi: 10.1109/TEPM.2005.862632.

[117] JEONG, K. M., SEON J., KYOUNG, K., KOH, C., CHOC, H. S. Development of PMP system for high speed measurement of solder paste volume on printed circuit boards. *Proc SPIE Optomecatronic Systems*. vol. 4564, no. 2001, pp. 250-259, 2001.

[118] GHIGLIA, D.C., PRITT, M.D. *Two-dimensional phase unwraping: Theory, algorithms and software*. first. New York: Wiley-intersience, 1998. ISBN 0-471-24935.

[119] BIOUCAS- DIAS, J., VALADAO, G. Phase Unwrapping via Graph Cuts. *IEEE Transactions Image Processing*. Vol.16, Issue.3, pp.698, 2007, ISSN: 10577149.

[120] HANI, A.F.M., KHOIRUDDIN, A.A., WALTER, N., FAYE, I. Wavelet analysis for shadow detection in Fringe Projection Profilometry. *Industrial Electronics and Applications (ISIEA), 2012 IEEE Symposium on*. vol., no., pp.336,340, 23-26 Sept. 2012 doi: 10.1109/ISIEA.2012.6496655.

[121] ZHANG , L., HE, X. Fake Shadow Detection Based on SIFT Features Matching. *Information Engineering (ICIE), 2010 WASE International Conference on*, vol.1, no., pp.216,220, 14-15 Aug. 2010, ISBN 978-1-4244-7506-3.

[122] WANG, Y., TANG, M., ZHU, G. An Improved Cast Shadow Detection Method with Edge Refinement. *Intelligent Systems Design and Applications, 2006. ISDA '06. Sixth International Conference on*, vol.2, no., pp.794,799, 16-18 Oct. 2006, doi: 10.1109/ISDA.2006.253714.

[123] HUANG, Ch-H., WU, R-Ch. An Online Learning Method for Shadow Detection. *In 2010 Fourth Pacific-Rim Symposium on Image and Video Technology. Singapore*. vol., no., pp.145,150, 14-17 Nov. 2010 doi: 10.1109/PSIVT.2010.31.

[124] SANIN, A., SANDERSON, C., LOVELL, B. C. Shadow detection: A survey and comparative evaluation of recent methods. *Pattern Recognition*. Vol. 45, issue 4, pp. 1684-1695, April 2012, ISSN 0031-3203.

[125] JINGKE, X. (2008) Outlier Detection Algorithms in Data Mining. *Intelligent Information Technology Application, IITA. Second International Symposium on* , vol.1, no., pp.94-97,2008,doi: 10.1109/IITA.2008.26.

[126] CHAVEZ, E. (2001) A subquadratic algorithm for cluster and outlier detection in massive metric data. *String Processing and Information Retrieval (SPIRE) , Proceedings.Eighth International Symposium on*, pp.46-58,2001,doi: 10.1109/SPIRE.2001.989736.

[127] GRUBBS, F. Procedures for Detecting Outlying Observations in Samples, *Technometrics*, vol. 11, no.1, pp. 1-21,1969.

[128] JOLLIFFE, I.T. Principal Component Analysis. *Springer Series in Statistics*. pp. 489, 2002, ISBN-10: 0387954422.

[129] STEFATO, G., HAMZA, A.B. Cluster PCA for outliers detection in high-dimensional data. *Systems, Man and Cybernetics, ISIC. IEEE International Conference on.* pp.3961-3966, 2007.

[130] SAHA B.N., RAY, N., HONG, Z. Snake Validation: A PCA-Based Outlier Detection Method. *Signal Processing Letters, IEEE*. vol.16, no.6, pp.549-552, 2009.

[131] RUSSO, R., MAXWELL, R. *Student's Guide to Analysis of Variance*. Routledge, 1999.

Own cited work

[132] BOLEČEK, L., ŘÍČNÝ, V., SLANINA, M. Modified Method for Optimization of Image Registration. *In TSP 2011*. první. Budapest: 2011. pp. 530-533. ISBN: 978-1-4577-1409- 2.

[133] BOLEČEK, L. *Zobrazování černobílých snímků v nepravých barvách*. Brno: Vysoké učení technické v Brně, Fakulta elektrotechniky a komunikačních technologií, 2009. 60 pp. Vedoucí semestrální práce prof. Ing. Václav Říčný, CSc

[134] BOLEČEK, L., ŘÍČNÝ, V. MATLAB Detection of shadow in Image of Profilometry. *In Technical Computing Prague 2011*. Praha: Humusoft s.r. o, 2011. pp. 22-30. ISBN: 978-80-7080-794- 1.

[135] KALLER, O., BOLEČEK, L., KRATOCHVÍL, T. Profilometry scaning for correction of 3D images depth map estimation. *In Proceedings of the 53rd International Symposium ELMAR- 2011*. Zadar, Croatia: ITG, Zagreb, 2011. s. 119-122. ISBN: 978-953-7044-12- 1.

[136] BOLEČEK, L., ŘÍČNÝ, V., SLANINA, M. Fast Method For Reconstruction of 3D Coordinates. *In TSP 2012*. první. Budapest: 2011.

[137] BOLEČEK, L., KALLER, O., ŘÍČNÝ, V. Influence of the Viewing Angle on the Spatial Perception for Various 3D Displays. *In Proceedings of 21st International Conference Radioelektronika 2012*. Brno: Vysoké učení technické v Brně, 2012.

[138] SLANINA, M., KRATOCHVÍL, T., BOLEČEK, L. ŘÍČNÝ, V.; KALLER, O.; POLÁK, L. Testing QoE in Different 3D HDTV Technologies. *Radioengineering*, 2012, vol. 22, no. 1, pp. 445-454. ISSN: 1210- 2512.

[139] BOLEČEK, L., ŘÍČNÝ, V. The Estimation of a Depth Map Using Spatial Continuity and Edges. *In 37th International Conference on Telecommunications and Signal Processing (TSP)*. first. 2013. pp. 51-54. ISBN: 978-1-4799-0403- 7.

[140] BOLEČEK, L., ŘÍČNÝ, V., KALLER, O. Statistical analysis of subjective tests results of the various 3D displays. *Slaboproudý obzor*. 2013, vol. 69, no. 4, pp. 11-17. ISSN: 0037- 668X.

[141] BOLEČEK, L., ŘÍČNÝ, V., SLANINA, M. 3D Reconstruction: Novel Method for Finding of Corresponding Points. *Radioengineering*, 2013, vol. 22, no. 1, pp. 82-91. ISSN: 1210- 2512.

LIST OF SYMBOLS, PHYSICAL CONSTANTS AND ABBREVIATIONS

$P(X, Y, Z)$ spatial point

X horizontal position in space, real word position

Y vertical position in space, real word position

Z depth position in space, real word position

$\mathbf{K}$ calibration matrix

f_c focus distance

(cc) principal point of the camera, represents optical center of the camera

ds radial distortion of the camera

$\mathbf{R}$ rotation matrix R

$\mathbf{Difference_{matrix}}$ matrix containing difference between potential position of the corresponding points in proposed method

$\mathbf{T}$ translation vector T

$\mathbf{E}$ essential matrix

$\mathbf{F}$ fundamental matrix

$\mathbf{H}$ homography matrix

B stereo base, distance between optical center of the cameras

$p_1\,(x_1, y_1)$ image point in first input image

$p_2\,(x_2, y_2)$ image point in first input image

x_{im} horizontal image coordinate in pixel

y_{im} vertical image coordinate in pixel

I image

λ unknow scale factor

$\mathbf{P_R}$ projective matrix of the right camera

$\mathbf{P_L}$ projective matrix of the left camera

SSIM Structural Similarity Index Measure

SA Spatial Activity

FA Frequency Activity

CC Correlation Coefficient

SD Standard deviation of the image

LE Local Entropy

LR Local Range

CO Contrast

K_r coefficient proposed for description image from view reliability of finding corresponding points

$p_{1,sel}\,(x_{1,sel}, y_{1,sel})$ image point in first input image

$p_{2,sel}\,(x_{2,sel}, y_{2,sel})$ image point in first input image

A_{im} area of the image

$depth$ depth, distance spatial points from the camera, equal to Z

$p_{1,SURF}$ image points found in I_1 by algorithm SURF

$p_{2,SURF}$ image points found in I_2 by algorithm SURF

m_x change of the horizontal position of the image of the spatial point P in two various image of the scene

m_y change of the vertical position of the image of the spatial point P in two various image of the scene

$p_{2,potential}(x_{2,potential}, y_{2,potential})$ potential position of the selected point in the I_2

$diff_{color}$ difference in color of the selected point in I_1 and of its potential position in I_2

R, G, B red, blue and green components of the RGB image

$diff_{i,j}$ differences between individual potential positions

α pitch error angle

β roll error angle

γ yaw error angle

ΔP overall error in the reconstruction

Z_{error} observed depth

Z_{real} real depth

Y_{error} observed vertical space position

Y_{real} real vertical space position

X_{error} observed horizontal space coordinate

X_{real} real horizontal space coordinate

ΔZ delta depth

ΔZ_C Calculated difference between Z_{error} and Z_{real}, this difference was experimentally determined and represents error caused by error in camera alignment.

ΔZ_T Theoretical difference between between Z_{error} and Z_{real}, this difference was obtained by proposed equations.

ΔX_C Calculated difference between X_{error} and X_{real}, this difference was experimentally determined and represents error caused by error in camera alignment.

ΔX_T Theoretical difference between between X_{error} and X_{real}, this difference was obtained by proposed equations.

ΔY_C Calculated difference between Y_{error} and Y_{real}, this difference was experimentally determined and represents error caused by error in camera alignment.

ΔY_T Theoretical difference between between Y_{error} and Y_{real}, this difference was obtained by proposed equations.

p_x horizontal paralax

$disp_{max}$ maximal disparity

win_{length} lebght of using window, number of using pixels

$grad_{weight}$ weight of the gradient

d horizontal shift of the image (disparity)

$diffG_{X,Y,Z}$ difference in gradient in individual color component

$differences$ overall difference for particular disparity and pixel

$depth$ depth map of the image

$Depth_{tolerance}$ tolerance of difference between depth

$depth_{border}$ depth on the border of zero region on the depth map

$region_{depth}$ length of the zero region

$image_{edge}$ edge representation

$depth_{profilometric}$ profilometric depth map

$depth_{stereo}$ stereo depth map

α_p angle of the projection

I_S image of the scene with projected pattern

I_R image of the scene with projected pattern

I_{shadow} shadow image

$I_{pattern}$ image pattern

I_{object} image with scene without projection

$I_{object_pattern}$ image with scene with projection

$depth_{final}$ final depth map

σ Standard Deviation

L_w viewing distance

α_w viewing angle

B_w viewing angle

S_i saturation of the color (i $\in \langle 1, 3 \rangle$)

B_i brightness of the color (i $\in \langle 1, 3 \rangle$)

G Results of the Grubbs test

x_r tested data point in Grubbs test, evaluation by certain respondent

3DTV television that conveys depth perception to the viewer by employing techniques such as stereoscopic display

SSD Sum of Squared Differences

SAD Sum of Absolute Differences

NCC Normalized Cross-corelation

SIFT Scale invariant Feature Transform

SURF Speed up robust Features

QoE Quality of experience

ITU International Telecommunication Union

DLT Direct Linear Transformation

H homography

LoG Laplacian of Gaussian

DoG Difference if Gaussian

SSIM Structural Similarity Index Measure

SA Spatial Activity

FA Frequency activity

CGRAD Cost from Gradient of Absolute Differences

H Hue, one of the component of the color space HSV

S Saturation, on of the component of the color space HSV

V Value, on of the component of the color space HSV

DLP digital data projector

FPP Fringe Pattern Profilometry

PSP phase shifting profilometry

FTP Fourier Transform Profilometry

LCD Liquid Crystal Display

HD High Definition

DTH Dynamic Head Tracking

PCA Principal Component Analysis

ANOVA Analysis of Variance