The Modelling of Stereoscopic 3D Scene Acquisition

Martin HASMANDA, Kamil RIHA

Dept. of Telecommunications, Brno University of Technology, Purkynova 118, 612 00 Brno, Czech Republic

hasmanda.martin@phd.feec.vutbr.cz, rihak@feec.vutbr.cz

Abstract. *The main goal of this work is to find a suitable method for calculating the best setting of a stereo pair of cameras that are viewing the scene to enable spatial imaging. The method is based on a geometric model of a stereo pair cameras currently used for the acquisition of 3D scenes. Based on selectable camera parameters and object positions in the scene, the resultant model allows calculating the parameters of the stereo pair of images that influence the quality of spatial imaging. For the purpose of presenting the properties of the model of a simple 3D scene, an interactive application was created that allows, in addition to setting the cameras and scene parameters and displaying the calculated parameters, also displaying the modelled scene using perspective views and the stereo pair modeled with the aid of anaglyphic images. The resulting modelling method can be used in practice to determine appropriate parameters of the camera configuration based on the known arrangement of the objects in the scene. Analogously, it can, for a given camera configuration, determine appropriate geometrical limits of arranging the objects in the scene being displayed. This method ensures that the resulting stereoscopic recording will be of good quality and observer-friendly.*

Keywords

Anaglyph, parallax, stereo base, stereo pair, stereoscopy.

1. Introduction

Stereoscopy has increasingly been used in the recent decades. This is due to the growth of the computing power of computers and to trying to obtain a view that corresponds to the perception a human has when observing an actual spatial scene. Currently, stereoscopy is used in many industries, e.g. in science, medicine, architecture, entertainment, and game industry, and also in other fields. For example, in robotics it is used for the orientation of autonomous robots in space and for object recognition [1]. It also finds wide application in medicine, displaying a stereoscopic view that improves the surgeon's perception when either teaching or performing the operation itself [2][3], in architecture for viewing interiors and exteriors of buildings, in product presentations in companies, etc. In recent years, making 3D Movies has also become very popular [4]. Stereoscopic videos are already well integrated into the video and audio compression standards MPEG [5]. Stereoscopy technology is also connected with viewing stereoscopic formats, in the anaglyphic form, which has been known since the 19th century. However, novel methods are constantly being developed that make watching more comfortable [6][7], leading to the more modern technologies using autostereoscopic or holographic units [8]. Another recent innovation is viewing stereoscopic formats on mobile phones with autostereoscopic displays [9].

People use a pair of eyes for the observation of the real-world and this vision is called binocular vision. In the first place, some of the terms need to be given that are currently used in binocular vision:

- *Disparity*: is the difference between the corresponding points from two images of the object images which are projected on the retina of the eye.

- *Parallax*: is the horizontal distance between the corresponding points from the left and the right image on a projector or display screen [5]. In other sources [10] the parallax is the angle between the lines of sight that leads to the disparity between the two retinal images. In this work the parallax will be used as defined in [5].

This paper describes the principle of calculating the camera separation (stereo base) for a pair of cameras with the same parameters in a parallel arrangement. Parallel arrangement means that the optical axes of both cameras are parallel.

Results of the research are realized via simulations to determine, how far the observed objects can be placed in the scene for the corresponding parallax in the image to be in standards that are appropriate for viewing stereoscopic videos [5].

When composing static stereoscopic images, the Bercovitz formula may be sufficient for good watching [11]. In the case of video sequences, the eyes get used to the small parallax value and the 3-D effect is not significant. It is therefore necessary to consider larger parallax values.

This paper aims to formulate specific parallax limits depending on the observed scene and the size of the stereo base of the camera system [12][13]. The parallax values should not reach large values, because too large values of positive and negative parallax can produce headaches and sore eyes

when watching stereoscopic video sequences, and should be reduced. They can also lead to an enhanced impact of the cross-talk [5] in a composite image and post processing is necessary such as that described in [14]. Using the negative parallax very often is not recommended either, because it increases the possibility of conflict with the stereoscopic image boundary [4], and the eyes are more strained than in the case of positive parallax.

Further it is necessary to prevent the appearance of the divergent parallax and the vertical parallax, discussed below. The following text will define the parameters characterizing the stereo pair. These include:

- The maximum positive and maximum negative parallax in the image, after the adjustment of stereo pair base to the necessary camera parameters such as focal length or horizontal viewing angle, scene parameters (size of the projection plane) and the value of the stereo base.

- The calculation of the stereo base for the desired parallax values, and setting the scene with selected parameters.

For a simple verification of mutual relations the resulting values are shown graphically in the application in the form of anaglyphic view.

In part 2 the basic principles and conditions for the stereoscopic projections and stereoscopic acquisition techniques for the stereo pair will be described. Potential parallax types will then be described that may occur in correctly composing a stereo pair and which should not occur. The main part describes the derivation of formulas for calculating the stereo base of the required positive parallax and the location of objects in the scene. At the end of this section the resulting relationships for computing the extremes of parallaxes will be presented. Equations are then derived for calculating camera parameters for perceiving depth in graphics applications after image adjustment.

The part 3 describes experimental results for real test stereo image data set. In this section, the difference between the real and modelled data is expressed.

The last section gives a simplified description of the application that was created based on calculations from the previous sections.

1.1 Stereoscopic Acquisition

Images taken by two cameras capturing the same scene are generally referred to as a stereo pair [15]. When composing images into the resulting stereoscopic format it is required that the two images of the stereo pair, from both the left and the right camera, be line-aligned. It is typical of the method described in this work that the images are mutually shifted by the value of the parallax in the image. This approach will be valid in the case that the cameras allow a displacement of the CCD in the horizontal direction of the

optical axes of the lenses and the resulting image will be given in the required image width W. Such a configuration is given in [17]. In the case that the acquisition of a scene is obtained by using cameras that have a fixed CCD position with respect to the optical axes of lenses, then further post- processing of the stereo pair is required. Such a situation will in the following be described as a stereo pair adjustment. The adjustment can be achieved by cropping the part of the left and the right image, which is not common (includes image information which is always located in only one of the images). Such trimming is given by the parallax value calculated for the object which after the acquisition is displayed on the screen.
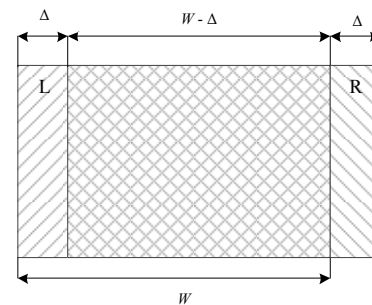

Fig. 1. The principle of the stereo pair acquisition with a parallel axis camera.

The original image width is given as W. The resulting composite image width W' will then be

$$W' = W - \Delta \tag{1}$$

where Δ denotes the distance between the corresponding points after the two images to be displayed on the image plane have overlapped. Parameter Δ will be referred to as image trimming. This procedure is shown in Fig. 1. The value Δ is one of the factors that need to be taken into consideration in order to be considered for the preserve as much usable image width in the resultant image as possible.

1.2 Parallax

In the human binocular visual system it is appropriate to mention two basic cases related to the subject matter of this work. The first of these is the effect of convergence, where the eyes are pointing towards the observed object. The second is the accommodation effect, where the lens changes its shape such that the image is projected on the retina of the eye. Accommodation corresponds to focusing the camera on the object [12][16].

In the human vision, the eyes converge on the object in the scene and accommodate to the observed object. When viewing the stereo image format, there is a difference in that the eyes can accommodate only to the projection plane (e.g. the monitor screen). Based on these principles, several situations can occur when watching a point in the projected scene. These approaches are described in [16]. For the sake of better understanding the principles of the projection of points

and for subsequent derivation of relationships these situations are briefly presented in Fig. 2 in the second section.
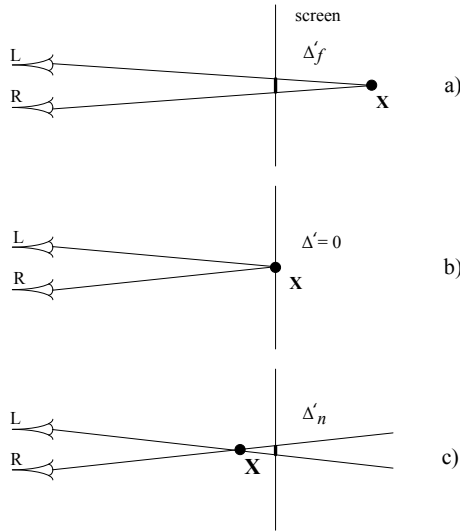

Fig. 2. The parallax types: a) positive parallax b) zero parallax c) negative parallax.

In Fig. 2, point $\mathbf{X}$ is the observed point in the scene. The parameters Δ'_n, Δ'_f and Δ' denote the values of the parallaxes in the image after adjustment. In the case of Fig. 2 a) point $\mathbf{X}$ is projected as if beyond the image plane, leading to the perception of depth in the scene. In the case of the zero parallax, point $\mathbf{X}$ is projected onto one point on the projection plane (as shown in Fig. 2 b) and the optical axes of the eyes are pointing towards to this point. The last state, shown in Fig. 2 c) occurs when the object appears as if it protruded from the monitor screen towards to the observer. The optical axes of the eyes in this case intersect in front of the screen.

All the situations from Fig. 2 commonly occur in the case of human binocular vision. When composing a stereo pair, a can also occur when the size of the positive parallax is larger than the eye separation Then the optical axes of the eyes diverge and this principle is known as divergent parallax [5]. In ordinary life this situation will never happen, and it needs to be avoided. In this article, the horizontal parallaxes will be discussed, i.e. shifting the image in the horizontal direction only. The vertical parallax is undesirable for composing and observing stereoscopic image formats [4].

2.   Calculation of the Stereo Base

Several approaches can be used to make stereo pairs of images of a scene. These approaches depend on the arrangement of cameras in the scene. More information about these methods can be found in [17][18]. One possibility, which is the closest to observation by humans, is characterized by the two cameras being rotated towards the observed object such that their optical axes intersect at one point, on which the cameras are focused. This approach is not suitable for making a stereo pair, because it gives rise to the undesirable vertical parallax in the image. A preferable solution is to use a system of cameras with parallel optical axes of the lenses.
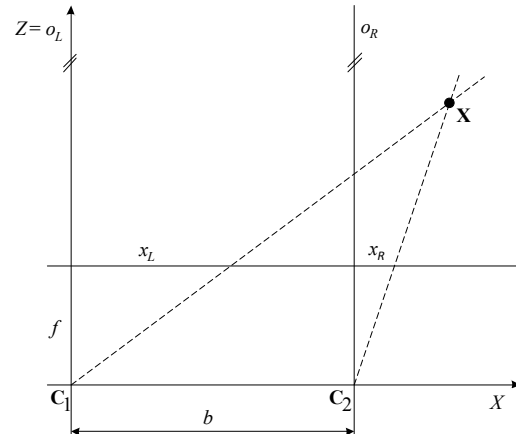

Fig. 3. Geometric situation for deriving the calculation of the stereo base in the parallel camera system.

The simplified principle of scene acquisition using a parallel camera pair is shown in Fig. 3, from which it is possible to derive a formula for calculating the stereo base, which describes the projection of the point $\mathbf{X} = (X, Y)^T$ into the projection planes of the two cameras with the optical axes in parallel. This configuration is based on two "pinhole" camera models. The pinhole camera model is given in [15] and describes perspective projection. The centers of the two cameras are labeled as $\mathbf{C}_1$ and $\mathbf{C}_2$. The distance between the centers of cameras is called the stereoscopic camera base. For the projection of the point $\mathbf{X}$ into the image planes of the cameras (when using two cameras with the same intrinsic parameters is assumed) could be determined on the basis of the similarity of triangles

$$X = x_L \frac{Z}{f}, \tag{2}$$

$$X = x_R \frac{Z}{f} + b. \tag{3}$$

The distance f on the Z axis determines the boundary of the objects that are displayed with a zero parallax. The distance from this point of this camera system is referred to as the zero boundary of the parallax. This is a boundary on which the object is located whose display the two cameras are adapted to. A precondition for a quality acquisition of the stereo pair is the choice of an appropriate distance between the camera centers, i.e. the stereo base b. In the ideal case, it should be equal to the eye separation of a human. This value is indicated in [5] as 65 mm in the adult person. In contrast to the human vision, the stereo base can be in stereoscopic acquisition be adapted to the scene characteristics (location of the objects in the scene and the actual physical width of the cameras as whole). From Fig. 3 and the corresponding relations (2) and (3) the equation for the stereo base b can be derived as

$$b = \frac{Z}{f} (x_L - x_R). \tag{4}$$

Here it is suitable to substitute $\Delta = x_L - x_R$ and then

$$b = \frac{1}{f} Z \Delta. \tag{5}$$

The variable Δ denotes the trimming of composite image between the right and the left view in the case of their overlap. Based on this parallax, a new image width in (1) will be calculated. The variable f determines the focal length of the image plane. This distance is related to the horizontal viewing angle of cameras and the actual width of the image plane.
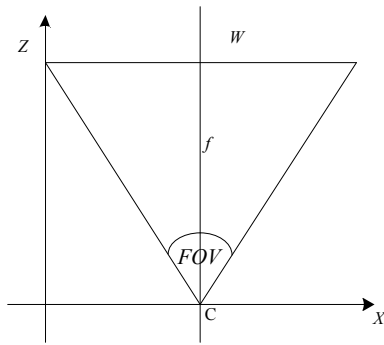

Fig. 4. The relationship of the focal distance due to the width of the screen.

In some cases, the value of camera focal length may not be given. This can be calculated using the horizontal viewing angle of the camera, denoted as FOV (Field of View), and the width of the projection plane from Fig. 4. With these parameters, the relationship for the focal length can be expressed as

$$f = \frac{W}{2\tan\left(\frac{FOV}{2}\right)}. \tag{6}$$

Then the stereo base from (6) can be modified to the final equation, which will be the width of the image plane W

$$b = 2\frac{\Delta}{W}Z\tan\left(\frac{FOV}{2}\right). \tag{7}$$

Equation (7) represents the basic formula for calculating the stereo base of the cameras with respect to the zero parallax boundary distance.

2.1 Calculating of the Parallax Limits in the Image

This chapter will describe the derivation of the formula for calculating the extreme values of the maximum positive and maximum negative parallax in the image.

Consider a parallel system of cameras shown in Fig. 5. It is further considered that the objects are placed behind the image plane in the Z direction. The farthest from the objects, $\mathbf{X}_F$, is located on the plane FP, the closest $\mathbf{X}_N$ on NP, and ZP is the plane on which an object $\mathbf{X}_Z$ is located, to which the cameras are adapted.
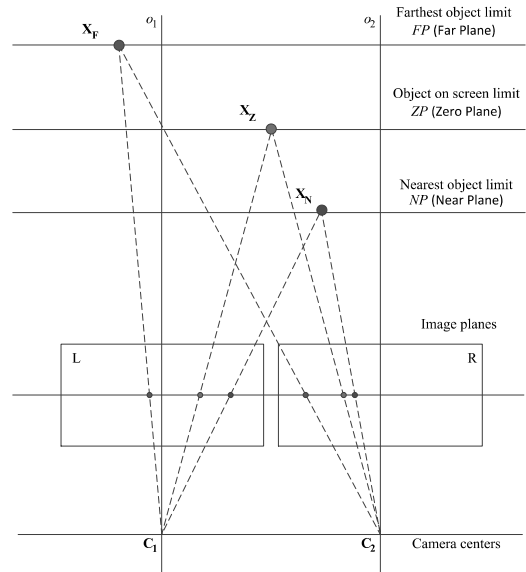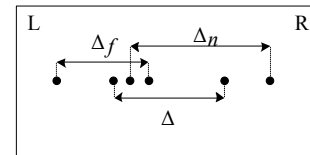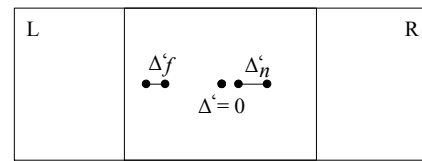

Fig. 5. The object locations in the scene.

Fig. 6. The parallax values during the scene acquisition.

Fig. 6 a) describes parallax variants Δ_n, Δ_f, Δ, which occur in the stereo acquisition in stereo pair images when they overlap without adjustment. In Fig. 6 b) the parallaxes Δ'_n, Δ'_f are then shown after adjustment via targeted shifting of the zero parallax plane. In Section 1.1 it was described that if any object was to be seen on the screen, this object must be shown with zero parallax value. Therefore, after the acquisition and the adjustment based on the perspective projection of two parallel cameras, two conditions can be established for the maximum values of both parallaxes (positive and negative) for points in the image planes. The general relationship from equation (7) can be simplified in order to find the limits wanted if the resultant width of the screen is known and if the size of the parallax is required in percentages of the width of the projection plane. Now it is appropriate to substitute

$$z = \frac{\Delta}{W} \tag{8}$$

where, for the sake of simplifying the calculation z gives the relative parallax in the image with respect to the composed image width. Using the substitution (8) the relationship between the stereo base and the scene settings is then expressed as

$$b = 2ZPz\tan\left(\frac{FOV}{2}\right). \tag{9}$$

From (9) equations for the limit parallax values can be derived. Considered in this article, will be z, f and n as the relative parallax values, and the post-acquisition parallax values f' and n'. Conversion to length units is possible using (8).

The relationship between the parallaxes is given by the perspective projection as:

$$f' \leq z \leq n' \tag{10}$$

where f' and n' are the relative values of parallaxes after the adjustment. The value z is the relative parallax on the image after image overlap prior to adjustment.

The second of the conditions is the relationship between the locations of objects

$$FP \geq ZP \geq NP \tag{11}$$

where FP, ZP and NP are the boundary planes.

2.2 The Determination of Maximum Parallax Values

This section describes how the positive and the negative parallax in the image can be determined from the calculated image trimming value and from the distances of objects in the scene.

After the acquisition of the stereoscopic image, the images are, according to section 1.1, shifted by the calculated value of the zero parallax. The zero parallax can be calculated from the stereo base from (9). This relationship can be used to calculate the parallax in the image at a distance Z from the system of camera lenses. To derive the resulting parallax after the adjustment these relationships can be used to determine the parallax between the projected points in both images. These corresponding relationships hold for the farthest object in the scene

$$b = 2FPf\tan\left(\frac{FOV}{2}\right) \tag{12}$$

and for the nearest object in the scene

$$b = 2NPn\tan\left(\frac{FOV}{2}\right) \tag{13}$$

where f and n are the parallaxes in the image before the adjustment.

The adjustment is given by (1) and describes the overlap of two images, which are mutually shifted by the parallax for the adjusted object. The value of the parallax Δ for this object before the adjustment can be calculated from the known stereo base from (7). By the principle of perspective projection the greatest parallax will be that of the object closest to the cameras system and, on the contrary, the smallest

parallax will be that of the farthest object. The parallaxes obtained after adjustment of the stereo pair in Fig. 6 b) and from (10) will be given by relations for the farthest object in the scene

$$f' = z - f \tag{14}$$

and for the nearest object in the scene

$$n' = z - n. \tag{15}$$

The parallax values before adjustment can be also expressed by (12) and (13) in the form

$$f = \frac{1}{FP} \frac{b}{2\tan\left(\frac{FOV}{2}\right)}, \tag{16}$$

$$n = \frac{1}{NP} \frac{b}{2\tan\left(\frac{FOV}{2}\right)}. \tag{17}$$

In the same way the value of the zero parallax in the image can be expressed from (9)

$$z = \frac{1}{ZP} \frac{b}{2\tan\left(\frac{FOV}{2}\right)}. \tag{18}$$

Relation (15) can be modified by substituting (17) and (18) as

$$n' = \frac{b}{2ZP\tan\left(\frac{FOV}{2}\right)}\left(\frac{ZP - NP}{NP}\right). \tag{19}$$

Then the resulting maximal negative parallax can be calculated from relations (15) and (17) as

$$n' = \frac{ZP - NP}{NP}z. \tag{20}$$

The distance NP must not be equal to zero, which also follows from the physical principles. If NP will be approaching zero, the resulting value of the negative parallax will approach infinity.

Similarly, the equation for the maximum positive parallax in the image can be derived from (16) and (18)

$$f' = \frac{FP - ZP}{FP}z. \tag{21}$$

The resulting relations (20) and (21) show, that the values of the positive and the negative parallax can be calculated from the knowledge of the layout of objects in a scene and from the known value of the parallax in the image.

2.3 The "Pop-Out" Distance

From the known observer distance from the scene and from the negative parallax value the approximate distance can be determined, at which the optical axes of eyes will converge in front of the screen. The distance of the projected object from the scene d is given by (22) as shown in Fig. 7.
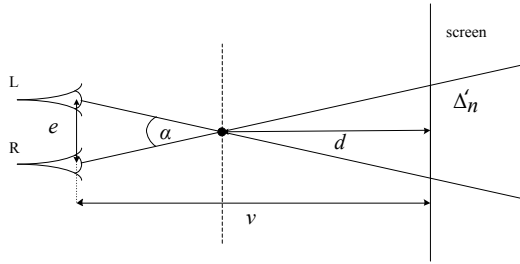
Fig. 7. The "pop-out" distance.

The parameter v is the distance of observer from the screen, and Δ_n' is the size of the maximum negative parallax. Then the distance d, at which the image will be projected in front of the screen, can be calculated as

$$d = \frac{v}{\left(\frac{e}{\Delta_n'}\right) + 1}, \qquad (22)$$

and α is the convergence angle of the eyes toward the object; it will decrease with increasing distance from the scene observer

$$\alpha = 2\tan^{-1}\left(\frac{\frac{e}{2}}{v - d}\right). \qquad (23)$$

2.4 Correction of Stereo Pair for Digital Processing

In the case of making an animated video or graphics applications the problem with the adjustment of the stereo pair as given in Section 1.1 can be solved by increasing the image width and the value of the field of view angle. When using the OpenGL or Direct3D interface, it is possible to set up any field of view angle and any desired visible area of the cameras. To find the appropriate relationships it is therefore sufficient to derive from the equation for the stereo base (7) a new value of FOV based on the parallax in image (25). Then the new value of the viewing angle FOV' can be calculated using the extended screen width W' from (6). The value of W' will be determined by extending the resulting W by the parallax in image according to (24)

$$W' = W + \Delta = W(1 + p). \qquad (24)$$

The resulting image width is given by (1) in the desired size and there would be no reduction of the resulting image size. The relevant horizontal viewing angle will be given as

$$FOV' = 2\tan^{-1}\left(\frac{W'b}{2\Delta Z}\right). \qquad (25)$$

This procedure is also used in the final application for displaying the results.

3. Experimental Results

3.1 Configuration of the Experiment

In this chapter, the tested dependence of the results on camera type and the accuracy of the method are described. Two common web cameras were used:

- **Camera1**: A4Tech PK-750MJ $f = 5.5$ mm, sensor CMOS $1/4"$ 3.2 mm (ca. 36.24 °);

- **Camera2** $C2$: Logitech QuickCam Pro4000 $EFL = 4.5$ mm, $FOV = 42$ °

These cheap web cams (although they were manufactured by the same manufacturer) can have different intrinsic parameters. Moreover, the camera case shape makes a precise CCD or CMOS sensor localisation difficult. The equations mentioned in Section 2 were derived for an exact parallel stereoscopic camera configuration. This problem was solved in the test data acquisition - it was performed by only one camera shifted by a stereo basis constant b.

The measurement was performed for very small distances between the objects and the camera. This small distance ensures a higher parallax and therefore better measurability of the resulting parallax.

A special scene configuration was prepared for testing the method. Cubic objects (a-sized cubes) are simple enough for edge detection and disparity estimation from test images (see Fig. 8). The sub-pixel edge detection method was used for accurate horizontal parallax measurement in the image.
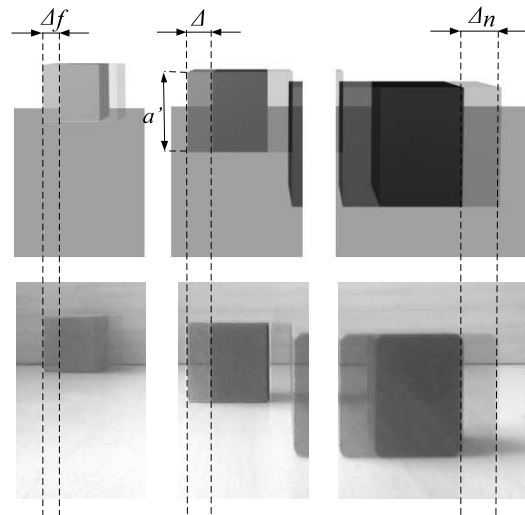

Fig. 8. The principle of measurements on the real and the synthetic scene.

Sub-pixel edge detection

In the first step, the pixel coordinate of the edge is estimated as a maximal value of the first derivative in the horizontal direction in the neighbourhood of the object edge. The second step is the polynomial approximation (2nd order) of the gradient values for the pixel positions around the detected (maximal) one. More details of this method can be found in [19].

FOV verification

The above methods can be used only if the value of FOV is known. Because of inaccurate data in the camera

data sheet, the real FOV was measured by the following procedure. 10 images were acquired by the camera from the same perspective, and object widths a'_i were measured (in pixel units). For error reduction, the average value $\overline{a'}$ was calculated. Using $\overline{a'}$ the pixel size in metric units can be calculated as

$$pxl = \frac{a}{\overline{a'}}. \tag{26}$$

The pixel sizes calculated for the two cameras were $pxl_{C1} = 0.43$ mm a $pxl_{C2} = 0.42$ mm. (For the real FOV measurement) To calculate the actual FOV value the real object (the a-sized cube) was placed at a distance Z in front of the camera. The equation for the calculation of the actual value of FOV is

$$FOV = 2\,tan^{-1}\left(\frac{imageResX\ pxl}{2\ z}\right) \tag{27}$$

where *imageResX* is the image resolution in the x direction in pixels.

The FOV values for both cameras was calculated as $FOV'_{C1} = 43.05°$ and $FOV'_{C2} = 41.78°$. As can be seen, there are significant differences between the real and data sheet values. Therefore, the calculated values of FOV were used in the following measurement.

3.2 Experiment with a Camera Basis Estimation

For this experiment, the object was placed at the distance $ZP = 350$ mm and the modelled values were calculated using equations (9), (12) and (13). The camera basis was changed in the range 0-100 mm and the values of Δ_i were measured in the image. The output values for both cameras were compared with the mathematical model Δ_{t_i} from (9) and plotted on the graph displayed in Fig. 9. The measurement error $\delta\Delta$ was calculated from the modelled and the measured values as $\delta\Delta_i = \Delta_{t_i} - \Delta_i$. Average measurement error $\overline{\delta\Delta}$ was $\overline{\delta\Delta_{C1}} = 1.63$ pxl for camera no. 1 and $\overline{\delta\Delta_{C2}} = 2.34$ pxl for camera no. 2.
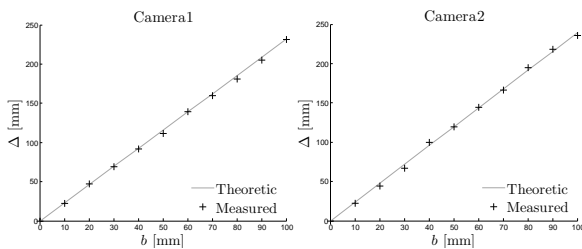

Fig. 9. Modelled and measured values of Δ in dependence on the size of stereo base.

The average measurement error of the positive parallax in metric units is then $\overline{\Delta z_{C1}} = 0.71$ mm and $\overline{\Delta z_{C2}} = 0.98$ mm.

3.3 Experiment with Disparity Estimation

The modelling values for this chapter were calculated using (20) and (21). Two objects of the same size were placed in the scene. The first object defines the distance $ZP = 350$ mm and the second one the variable distances NP and FP. The parallax values were calculated from acquired images for each step (10 mm). The stereo basis value for a given scene was chosen empirically as $b = 10$ mm. This value ensures that the projected objects will not collide with image borders for the chosen values of NP and FP. The value of z calculated from the given stereo base was $z_{C1} = 0.067$ for camera no. 1, and $z_{C2} = 0.1$ for camera no. 2.

3.3.1 Results for the NP

The NP distance is changing in the range $NP_i = 350 - 250$ mm in steps of 10 mm. For each step (image), the value $\Delta'_{n_i} = \Delta_i - \Delta_{n_i}$ was estimated. The values Δ'_{n_i} were then compared with the theoretical values $\Delta'_{n_{t_i}}$ using (20). These dependence values are plotted on the graphs in Fig. 10.
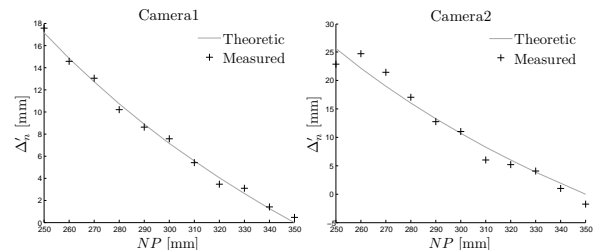

Fig. 10. Modelled and measured values of negative parallax in dependence on the NP.

The average measurement error of negative parallax estimation with NP changing was $\delta\Delta'_{n_i} = \Delta'_{n_i} - \Delta'_{n_{t_i}}$, $\overline{\delta\Delta'_{nC1}} = 0.32$ pxl and $\overline{\delta\Delta'_{nC2}} = 1.25$ pxl.

3.3.2 Results for the FP

The FP distance is changing in the range 350-450 mm in steps of 10 mm. For each step (image), the value $\Delta'_{f_i} = \Delta_i - \Delta_{f_i}$ was estimated. The values Δ'_{f_i} were then compared with the theoretical values $\Delta'_{f_{t_i}}$ using (21). These dependence values are plotted on the graphs in Fig. 11.
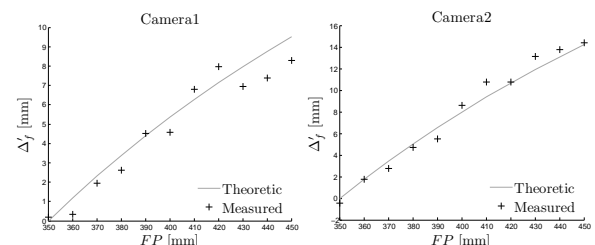

Fig. 11. Modelled and measured values of positive parallax in dependence on the FP.

The average measurement error of positive parallax estimation with FP changing was $\delta\Delta'_{f_i} = \Delta'_{f_i} - \Delta'_{f_{t_i}}$, $\overline{\delta\Delta'_{fC1}} = 0.62$ pxl and $\overline{\delta\Delta'_{fC1}} = 0.59$ pxl.

4. Software

The principles referred to above were modeled in an artificial 3-D scene using the relationships described for calculating the derivation of the relations between the distance of objects and the stereoscopic base. This application was written using the Microsoft Visual C# with Framework 4.0. It is available for downloading at [20]. The application for the display of the principles mentioned uses a 3D scene and their corresponding anaglyphic format, which is calculated on the basis of the parameters selected. The application can be used for monitoring and setting the parameters of a stereoscopic camera system with parallel optical axes.

In Fig. 12 the setting of an example model of the stereoscopic system is shown. The values of stereoscopic scene parameters and screen parameters are plotted here. The parameters of the cameras in the scene and of the observed scene itself are set in the left panel. The object and camera locations in the scene can be observed in the 3D window. The presentation of the final composition of the stereo pair on the basis of the calculated parameters can for visual verification of the calculations be followed on the anaglyphic image.
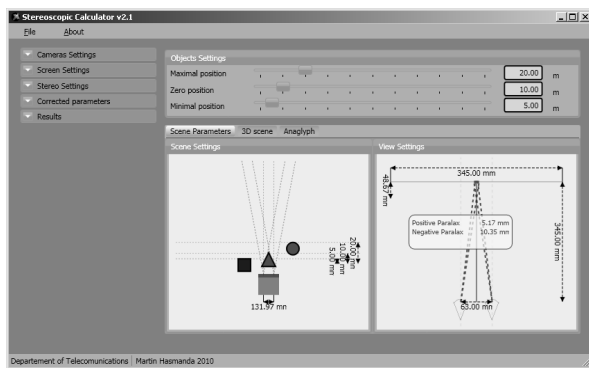

Fig. 12. The resulting application for verification of the computed equations.

5. Conclusion

The paper describes the derivation of the relations between the camera base and the distance of objects in the scene. The result is a mathematical definition of the boundary conditions for the adaptation of the stereo pair for its presentation on a suitable 3-D display or on a current display, using appropriate tools such as 3D glasses. Relationships between the stereo base and maximal positive and maximal negative parallax in the image, formed after the acquisition of the images captured by two cameras in a parallel optical axis arrangement, have been described and derived. These parallax values are essential to human perception and acquisition of the stereo pair. The parallax values should be small enough to avoid excessive stress on eyes due to their unnatural convergence and accommodation. The properties of stereoscopic perception can be corrected via targeted mutual shifting of the stereo pair by a suitable value of the parallax.

The results were implemented by way of an individual application that models the characteristics of the stereoscopic camera system and enables the determination of the appropriate parameters. The influence of these parameters on the display properties is simulated on computer-generated images created using the demo scene in Direct3D. The testing of presented mathematical model for real scene data was also performed. Real cameras do not have the same intrinsic parameters of production and their precise location in a parallel configuration to prevent vertical parallax can be problematic. In this case, the actual acquisition of the scene should be preceded by the calibration of the stereoscopic camera system and rectification [15] of the two stereo pair images. Experimental results for images 640 x 480 pxl in size showed that the average measurement error of disparity estimation with a variable camera base width was approximately 1 - 2 pxl. The average measurement error of disparity estimation with a constant camera base width and variable distance of the object from the camera was in most cases less than 1 pxl.

Acknowledgements

References

[1] CHIOU, R. Y., KWON, Y., TSENG, T.-L., KIZIRIAN, R., YANG, Y.-T. Real-time stereoscopic 3D for e-robotics learning. *Journal of Systemics, Cybernetics and Informatics*, 2011, vol. 9, no. 1, p. 8 - 13.

[2] JOHN, N. W. Using stereoscopy for medical virtual reality. *Medicine Meets Virtual Reality*. Amsterdam: IOS Press, 2002, p. 214 - 220.

[3] DEO, J. D., DE, S., TEJINDER, S. P. Using stereoscopy for medical virtual reality. In *Medicine Meets Virtual Reality 15*. Amsterdam, 2007, p. 115-120.

[4] MENDIBURU, B. *3D Movie Making : Stereoscopic Digital Cinema from Script to Screen*. Oxford: Focal Press, 2009.

[5] JAVIDI, B., OKANO, F. *Three-Dimensional Television, Video, and Display Technologies*. Berlin: Springer, 2002.

[6] IDESES, I., YAROSLAVSKY, L. New methods to produce high quality color anaglyphs for 3-D visualization. *Image Analysis And Recognition: Lecture Notes in Computer Science*, 2004, vol. 3212/2004, p. 273 - 280.

[7] MCALLISTER, D. F., ZHOU, Y., SULLIVAN, S. Methods for computing color anaglyphs. In *Proceedings of SPIE: Stereoscopic Displays and Applications XXI,* 2010, vol. 7524, p. 75240S-75240S-12.

[8] SON, J.-Y., JAVIDI, B., KWACK, K. -D. Methods for displaying three-dimensional images. In *Proceedings of the IEEE* 2006, vol. 94, no. 3, p. 502 - 523.

[9] FLACK, J., HARROLD, J., WOODGATE, G. J. A prototype 3D mobile phone equipped with a next generation autostereoscopic display. In *Conference on Stereoscopic Displays and Virtual Reality Systems XIV.* San Jose (USA), 2007, p. 12.

[10] MULLER, K, MERKLE, P., WIEGAND, T. 3-D video representation using depth maps. In *Proceedings of the IEEE*, 2011, vol. 99, no. 4,p. 643 - 656.

[11] BERCOVITZ, J. Image-side perspective and stereoscopy. *Proceedings of SPIE 3295,* 1998, p. 288 - 298.

[12] LAMBOOIJ, M., FORTUIN, M., IJSSELSTEIJN, W. Measuring visual fatigue and visual discomfort associated with 3-D displays. *Journal of the Society for Information Display*, 2010, vol.18, no. 11, p. 931 - 943.

[13] HOWARTH, P. A. Potential hazards of viewing 3-D stereoscopic television, cinema and computer games: a review. *Ophthalmic & Physiological Optics*, 2011, vol. 31, no. 2, p.111 - 122.

[14] KRUPEV, A. A., POPOVA, A. A. Ghosting reduction and estimation in anaglyph stereoscopic images. In *IEEE International Symposium on Signal Processing and Information Technology ISSPIT 2008.* Sarajevo (Bosnia and Herzegovina), 2008, p. 375 - 379.

[15] HARTLEY, R., ZISSERMAN, A. *Multiple View Geometry in Computer Vision.* 2nd ed. Cambridge (UK): Cambridge University Press, 2004.

[16] STEINMAN, S. B., STEINMAN, B. A., GARZIA, R. P. *Foundations of Binocular Vision: A Clinical Perspective.* New York (USA): McGraw-Hill, 2000, p. 357.

[17] KIM, H. D., SOHN, H. K. Depth adjustment for stereoscopic image using visual fatigue prediction and depth-based view synthesis. In *2010 IEEE International Conference on Multimedia and Expo ICME 2010.* Singapore, 2010, p. 956 - 961.

[18] WOODS, A., DOCHERTY, T., KOCH, R. Image distortions in stereoscopic video systems. In *Stereoscopic Displays and Applications IV.* Singapore, 1993, p. 36 - 48.

[19] HAGARA, M., KULLA, P. Edge detection with sub-pixel accuracy based on approximation of edge with Erf function. *Radioengineering*, 2011, vol. 20, no. 2, p. 516 - 524.

[20] HASMANDA M. *The Stereoscopic Calculator Software*. [Online] Cited 2010-30-06. Available at: http://www.utko.feec.vutbr.cz/~riha/software/StereoCalculator /StereoCalculator.exe.

About Authors . . .

Martin HASMANDA was born in 1985 in Krnov, Czech Republic. He received his M.Sc. degree in Electrical Engineering in 2010 from Brno University of Technology, Czech Republic. His research interests include 3D reconstruction technologies and digital image processing.

Kamil RIHA was born in 1978 in Nove Mesto na Morave, Czech Republic. He received his M.Sc. degree in Electronics & Communication in 2003 and the Ph.D. degree in 3D Scene Acquisition for Auto-Stereoscopic Display in 2007. Presently, he is employed at Brno University of Technology, Faculty of Electrical Engineering, Department of Telecommunications, as the academic employee (since 2006). His research interests include in particular areas of digital image and video processing.