Planning of a multiple sensor system for human activities space – aspects of iso-disparity surface

Jiandan Chen[*a], Siamak Khatibi[a], Jenny Wirandi[b] and Wlodek Kulesza[a]

[a]Blekinge Institute of Technology, SE-371 79 Karlskrona, Sweden,
[b]University of Kalmar, SE-391 82 Kalmar, Sweden

ABSTRACT

The Intelligent Vision Agent System, IVAS, is a system for automatic target detection, identification and information processing for use in human activities surveillance. This system consists of multiple sensors, and with control of their deployment and autonomous servo. Finding the optimal configuration for these sensors in order to capture the target objects and their environment to a required specification is a crucial problem. With a stereo pair of sensors, the 3D space can be discretized by an iso-disparity surface, and the depth reconstruction accuracy of the space is closely related to the iso-disparity curve positions. This paper presents a method to enable planning the position of these multiple stereo sensors in indoor environments. The proposed method is a mathematical geometry model, used to analyze the iso-disparity surface. We will show that the distribution of the iso-disparity surface and the depth reconstruction accuracy are controllable by the parameters of such model. This model can be used to dynamically adjust the positions, poses and baselines lengths of multiple stereo pairs of cameras in 3D space in order to get sufficient visibility and accuracy for surveillance tracking and 3D reconstruction. We implement the model and present uncertainty maps of depth reconstruction calculated while varying the baseline length, focal length, stereo convergence angle and sensor pixel length. The results of these experiments show how the depth reconstruction uncertainty depends on stereo pair's baseline length, zooming and sensor physical properties.

Keywords: iso-disparity surface, multiple surveillance sensors, stereo vision, sensor configuration, 3D reconstruction

1. INTRODUCTION

The Human ability to process visual information may be extended with the help of advanced technologies. The Intelligent Vision Agent System, IVAS, is one such high-performance autonomous distributed vision and information processing system. The system involves collecting data with different levels of speed and accuracy, in order to reconstruct 3D information for security, health care, and surveillance applications. The system is able to focus on the important and informative parts of a visual scene by dynamically controlling the pan-tilt-zoom of a stereo pair. For such a system, the critical problem is to find the optimal configurations of sensors and to gain the required level of reconstruction accuracy. The stereo pair camera's profile, such as baselines, convergence angles, pixel size, and focal lengths, are the most influential factors in determining the accuracy of 3D reconstruction. The effect of these parameters on the 3D reconstruction can be analyzed from the shapes and positions of the iso-disparity surfaces in 3D space.

The shape of the iso-disparity surfaces for general stereo configurations was first studied by Pollefeys et al., [1]. They described how the iso-disparity surfaces characterize the uncertainty and discretization in stereo reconstruction. A qualitative analysis of iso-disparity curves is also given. Although the geometry of these surfaces is well known in the standard stereo case, i.e. the front-parallel camera, there is a lack of analysis for the general stereo configuration. Also, the quantitative analysis of the iso-disparity surfaces has not often been studied.

This paper presents the iso-disparity surface geometry model, which is important in optimizing the stereo pair's configuration and in precisely anticipating the depth reconstruction accuracy. In addition, the model can also be used to make assumptions for many stereo algorithms, since these algorithms make hypothesis relying on the disparity range, i.e. the matching algorithm, [2]. In active vision, by estimating the disparity, control of the stereo convergence angle has already been introduced, [3]. The iso-disparity geometry model in active vision can help to select the disparity range according to surface geometry of the target.

---

[*] jian.d.chen@bth.se; phone +46 480 446774; fax +46 480 446330

Consideration of the iso-disparity when calculating the reconstruction uncertainty has been discussed by Völpel and Theimer, [4], where the disparity is considered in the *x*- and *y*- direction without using the epipolar geometry approach. In this paper, the disparity is defined along the epipolar line and the reconstruction uncertainty problem is solved by the iso-disparity geometry equations. This can be applied in more general cases.

A simple factor which helps to control the depth reconstruction accuracy is introduced in [5]. This paper improves the analysis of depth reconstruction accuracy into the stereo pair's *F*ield of *V*iew, FoV. This paper also considers the adjustment of stereo baseline for one stereo pair with a view to improving the depth reconstruction accuracy. The dynamic adjustment of stereo baseline for a parallel stereo pair was introduced in [6].

2. DEFINITIONS AND PROBLEM FORMULATION

The depth reconstruction accuracy can be controlled by adjusting the intervals of iso-disparity surface. This gives the possibility of planning a multiple sensor system, which can be implemented to observe human activities in 3D space with the required depth accuracy.

2.1 Definitions

Similar to the human eyes, stereo vision observes the world from different points of view. Two images are needed which are the fused to obtain a depth perception of the world. Any point in the world scene is captured in these two images as corresponding points which lie on the corresponding epipolar lines. It is necessary to define three terms related to the depth reconstruction: disparity, depth reconstruction uncertainty and depth reconstruction accuracy.

Disparity in this paper refers to the displacement of corresponding points along the corresponding epipolar lines for a common scene point, [1]. In the case where epipolar lines are horizontal the disparity is measured directly from the difference of the corresponding points' coordinates. The inverse projection of all possible image points with the same disparity will reconstruct the iso-disparity surfaces in 3D space.

Depth reconstruction uncertainty is defined as the intervals between discrete iso-disparity surfaces due to the discrete sensor. The depth reconstruction accuracy is the inverse of the depth reconstruction uncertainty.

2.2. Problem Statement and Main Contributions

The depth reconstruction may be calculated from a stereo pair with an accuracy determined by the system configuration. The system configuration is defined by sensor resolution (pixel size), focal lengths, baseline length and convergence angle. To get a more accurate depth reconstruction, the stereo configuration can be adjusted within its limits. The proposed models are limited to a general parallel stereo pair with zooming and the convergence stereo pair.

The main contributions of the paper can be summarized as follows:

- To model the iso-disparity surfaces for a general parallel stereo pair with zooming, as function of the baseline length, focal lengths and sensor pixel size.
- To model the iso-disparity surfaces for a general convergence stereo pair as function of the baseline length, convergence angle, focal length and sensor pixel size.
- Using iso-disparity geometry surfaces to quantitatively analyze the depth reconstruction accuracy.

3. PROBLEM ANALYSIS

The iso-disparity surfaces of a stereo pair may be simulated using synthetic methods. However such simulation is time consuming, and for planning real-time multi sensor system an easy mathematical model of the iso-disparity surfaces is needed.

There are two configurations for a stereo pair in common use. The first one is a parallel stereo pair in which the optical axes of the cameras are parallel. The cameras may have the same focal lengths or their focal lengths may be different, e.g. to get better reconstruction accuracy of a target placed at the boundaries of cameras' field of view. The second common configuration is a convergence stereo pair, where the optical axes cross at a fixation point. The simple mathematical models of iso-disparity surfaces for these configurations are analyzed in this chapter.

3.1 The iso-disparity surface of a parallel stereo pair

From the geometry of a parallel stereo pair, two cameras with parallel optical axes, with different focal lengths for left and right camera, f_L and f_R respectively, the iso-disparity plane for disparity $n\Delta D$ can be defined as:

$$z(x,n) = \frac{f_L - f_R}{n\Delta D} x + \frac{B}{2n\Delta D}(f_L + f_R) \tag{1}$$

where B is baseline length, n is integer number , ΔD is the disparity resolution. The planes are shown as the thin green lines in Fig. 1(a) and Fig. 1(c).

All the iso-disparity planes intersect with the xy-plane (the stereo pair baseline is a part of x-axis), and converge to the straight line:

$$x = \frac{B}{2}\frac{f_L + f_R}{f_R - f_L}, \quad z = 0 \qquad (f_L \neq f_R) \tag{2}$$

It is clear from equation (1) that when the focal lengths are equal $f_L=f_R=f$, z is independent of x and the iso-disparity planes are parallel to xy-plane, see the thin green lines in Fig. 1(b).

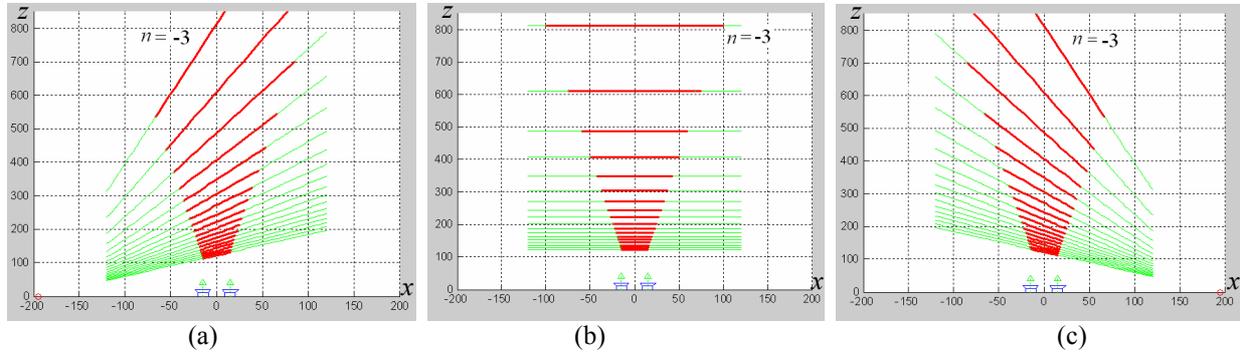From the inverse projection of image points and the triangulation method, using the Epipolar Geometry

Fig. 1. Iso-disparity planes for parallel stereo pair from synthetic (bold red line) and a plot of the mathematical model from equation (1). The lines are plotted with steps of 10 pixels. (a) Cameras with different focal lengths, $f_L = 3.5$ cm, $f_R = 3.0$ cm for left and right camera respectively. The convergence point is (-195 cm, 0) on xz-plane. (b) Cameras with the same focal length of 3.25 cm. (c) Cameras with different focal lengths $f_L = 3.0$ cm, and $f_R = 3.5$ cm for left and right cameras respectively, the convergence point is (195 cm, 0) on xz-plane.

Toolbox, [7], we can get the synthetic iso-disparity surfaces. Fig. 1 shows the synthetic disparity surfaces (the bold red lines) and the plots from equation (1) (the thin green lines). Here the baseline length B is 30 cm and the disparity resolution $\Delta D=0.04$ cm, or ten sensor pixel lengths where $p=0.004$ cm. Fig. 1(a) and Fig. 1(c) are plotted for the parallel stereo pair with different focal lengths. The parallel iso-disparity planes for parallel stereo pairs with same focal lengths are shown in Fig. 1(b). The synthetic simulation and calculating the equation give the same results.

3.2 The iso-disparity surface of convergence stereo pair

Let us consider two cameras with a convergence angle α_c, where $\alpha_{cL0}=\alpha_{cR0}=\alpha_c$ for the left and right camera respectively, with the angles rotated inwards to achieve a fixation point FP_0, as in Fig. 2. If the point TP_0 lies on the baseline's axis of symmetry, then the angles, (ψ_{L0}, ψ_{R0}), are the angles between the visual lines and a line perpendicular to the baseline. The zero disparity circle is defined by the fixation point and the left and right camera position points C_L and C_R. This circle is known as Vieth-Müller circle, and is a projection of the horopter, [8].

The iso-disparity surface is a cylinder whose cross section on the xz-plane is a conic that passes through both the centers of projection C_L, and C_R, and the point M_∞. M_∞ is a point imagined at infinity in both images, which can be obtained from the intersection of the normals to the optical axes, going through the projection centrals, [1]. It is possible to prove that for the case when $\alpha_{cL0}=\alpha_{cR0}=\alpha_c$, the conic is an ellipse. We need to define the ellipse's five degrees of freedom. Three of these are determined by the points C_L, C_R and M_∞. One of the two remaining degrees is related to the
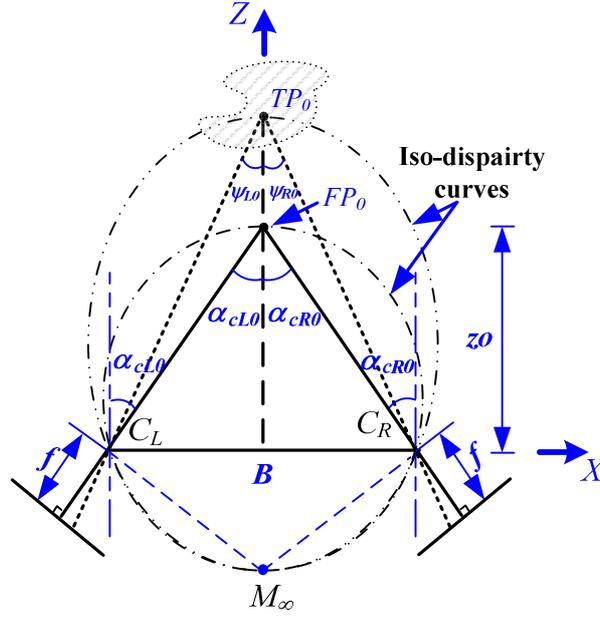
Fig. 2. An example of the iso-disparity curves for the convergence stereo pair in the plane defined by the cameras optical axes. z_0 is the distance from the fixation point to the baseline, f is the focal length.

point TP_0 with the disparity $n\Delta D$. The relationship between disparity $n\Delta D$ and focal lengths f_L and f_R for left and right cameras respectively, is the last required degree of freedom. If the disparity $n\Delta D$ and focal lengths f_L and f_R are known, the unique ellipse can be determined.

The iso-disparity surface of discrete disparity $n\Delta D$ for a convergence stereo pair (C_L, C_R) with the same focal length f and same convergence angles $\alpha_{cL0}=\alpha_{cR0}=\alpha_c$, describes a cylinder, the ellipses being cross sections of this cylinder on the xz-plane with centres in $0_e(x_{0e}(n), z_{0e}(n))$:

$$\frac{(x - x_{0e}(n))^2}{a^2} + \frac{(z - z_{0e}(n))^2}{b^2} = 1 \qquad (3)$$

For the chosen coordinates $x_{0e}=0$ and $z_{0e}=b-B\tan\alpha_c/2$:

$$\frac{x^2}{a^2} + \frac{\left(z - \left(b - \frac{B}{2}\tan\alpha_c\right)\right)^2}{b^2} = 1 \qquad (4)$$

where B is the baseline length and α_c is the stereo convergence angle. The ellipse half-axis along the z-axis $b(n,\Delta D,B,f,\alpha_c)$ depends on the discrete disparity $n\Delta D$, baseline length B, focus length f and convergence angle α_c and is described as:

$$b = \frac{\dfrac{B}{2}}{\sin 2\alpha_c - \dfrac{n\Delta D}{f}\cos^2\alpha_c} \qquad (5)$$

The ellipse half-axis along the x-axis $a(n,\Delta D,B,f,\alpha_c)$ can be found from the relationship:

$$\left(\frac{b}{a}\right)^2 = \tan\alpha_c \frac{1 + \dfrac{n\Delta D}{2f}\tan\alpha_c}{\tan\alpha_c - \dfrac{n\Delta D}{2f}} = \frac{\tan\alpha_c}{\tan\psi_c} \qquad (6)$$

where $\psi_c = \psi_{L0} = \psi_{R0}$.

The result of the synthetic stereo pair simulation with a baseline length 50 cm, focal lengths 2.5 cm and disparity resolution ΔD=0.04 cm is shown in Fig. 3. The synthetic iso-disparity surfaces (bold red lines) can be compared with ellipses from the equations (4)-(6) (thin green lines). Fig. 3 shows both the synthetic iso-disparity surfaces and the ellipses from the equation (4) in 3D space, in perspective view in Fig. 3(a) and top view in Fig. 3(b). Both results match each other perfectly.

4. APPROACH

Since the gaps between iso-disparity surfaces represent the discretization uncertainty in 3D space, we can generate a 3D depth reconstruction uncertainty map of a particular stereo pair's configuration using the iso-disparity surface geometry equations (4)-(6). Also, it is possible to generate such a map in 2D on the optical axes plane. This map can be used for the optimization of the stereo setup configuration. Generation of the 2D uncertainty map for a stereo pair configuration can be done in the following three steps.

Firstly, the plane has to be covered by the stereo pair's FoV, [9]. The area is sampled using small grids covered by the stereo pair.

Secondly, an iso-disparity curve on the optical axes plane should be calculated, passing through each grid point. Knowing that the curve will have a canonical shape, five points are needed. Two of these points can be the grid point and its symmetrical point, with respect to the symmetry axis of the baseline. The three others points are C_L, C_R and M_∞. For a convergent stereo pair, the ellipse axes a and b can be found using the ellipse fitting algorithm, [10]. Then using equation (5), the two closest ellipses with discrete disparity values $n\Delta D$ and $(n+1)\Delta D$ respectively, can be found, where the disparity resolution ΔD is one sensor pixel length.

Finally, the depth reconstruction uncertainty can be calculated as the interval between the iso-disparity surfaces, with the disparity values, $n\Delta D$ and $(n+1)\Delta D$ as the distance between the intersections of these two iso-disparity surfaces, and the line through the grid point and M_∞.

(a)   (b)
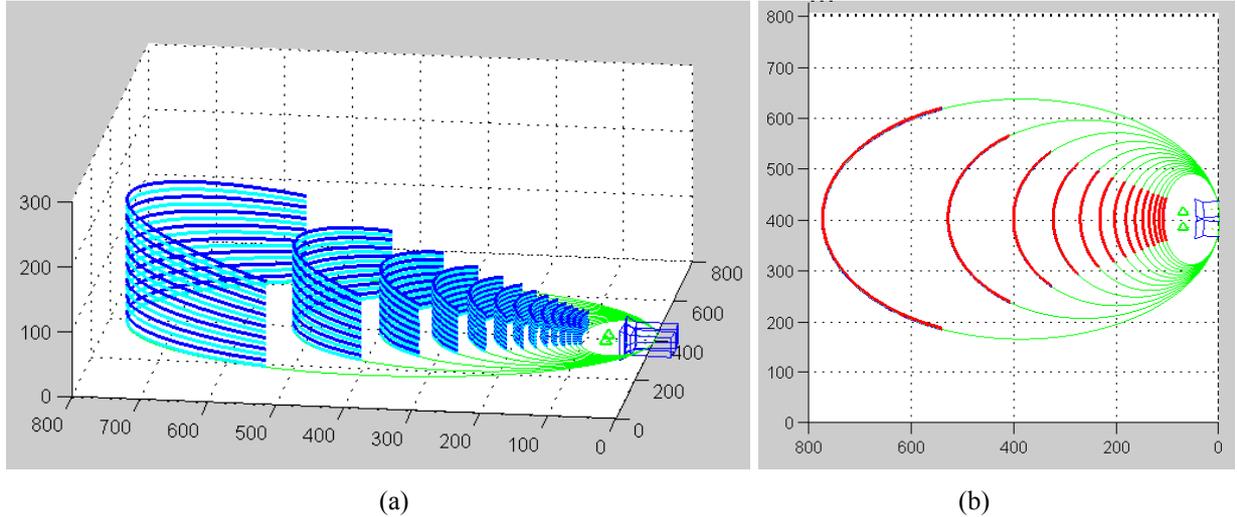
Fig. 3. Simulation results of iso-disparity surfaces for a stereo pair from the synthetic model (bold red line) and mathematical model from equation (4) (green line) with convergence angle, α_c=4°, the baseline length B=40 cm, the focal lengths f=2.5 cm and disparity resolution ΔD=0.04 cm (a) perspective view, and (b) top view.

5. RESULTS

The simulations presented were performed in MATLAB 7.0, and cover a rectangular area of (800 cm × 800 cm). This case study illustrates how depth reconstruction uncertainty in stereo coverage varies with the target distance for a given stereo baseline length, focal length, and sensor pixel length. The results are presented in Fig. 4, where the cameras optical axes are in the *xz*-plane. The depth reconstruction uncertainty is specified by the positive *y*-axis of the coordinate. However, this uncertainty analysis shows only the area covered by the stereo pair's FoV. To scale the uncertainty on the optical axes plane, a colour map is used. The lowest uncertainty is indicated by the blue colour and the highest uncertainty by the red colour. In order to increase the readability of the iso-disparity curves, the contour is plotted with ten pixel lengths disparity resolution. The map of the iso-disparity curves is generated with baseline length 40 cm, focal length 3.5 cm and pixel length $p=0.004$ cm, stereo convergence angle, $\alpha_c=4°$ and the FoV is approximately 54°. This case study proves that the depth reconstruction uncertainty increases as the distance to the target increases.

To show the discrete properties of depth reconstruction uncertainty, the map of the iso-disparity curves with suitable baseline length and pixel length is shown in the Fig. 5. The figure shows only half of FoV, with a cross section along the ellipses' axes perpendicular to the baseline. The discretization step increases with the target distance.

An exact illustration as to how the depth reconstruction uncertainty varies with the baseline lengths, focal lengths, sensor pixel length and stereo convergence angle, is shown in Fig. 6 and Fig. 7. Fig. 6(a) shows that the relative depth reconstruction uncertainty, relative to the target distance, decreases when the baseline length increases. The relative uncertainty decreases slowly for a baseline above about 40 cm. Its minimum value tends to be constantly between 0.5% and 1.5% for target distances of 200 cm and 800 cm respectively. At the same time, for a baseline of about 10 cm, the uncertainty varies between 10% and 2.5% for the respective target distances.

The change of the relative uncertainty versus the focal length is similar to that of the baseline length; see Fig. 6(b). For a focal length of longer than 3.5 cm the increase of the uncertainty is relatively slow. Its minimum tends to be constantly between 1.5% and 0.4% for target distances of 200 cm and 800 cm respectively. Meanwhile, for a focal length of 1 cm, the uncertainty varies between about 9% and 2% for the respective target distances.

Furthermore, Fig. 7(a) illustrates the linearly relation of the relative uncertainty and the sensor pixel length. Within the range from 0.001 cm to 0.006 cm, the relative uncertainty varies from 0.2% to 3.5% and depends also on the target distance.

Fig. 7(b) shows that the stereo convergence angle has a slight influence on the uncertainty but this also depends on the target distance.
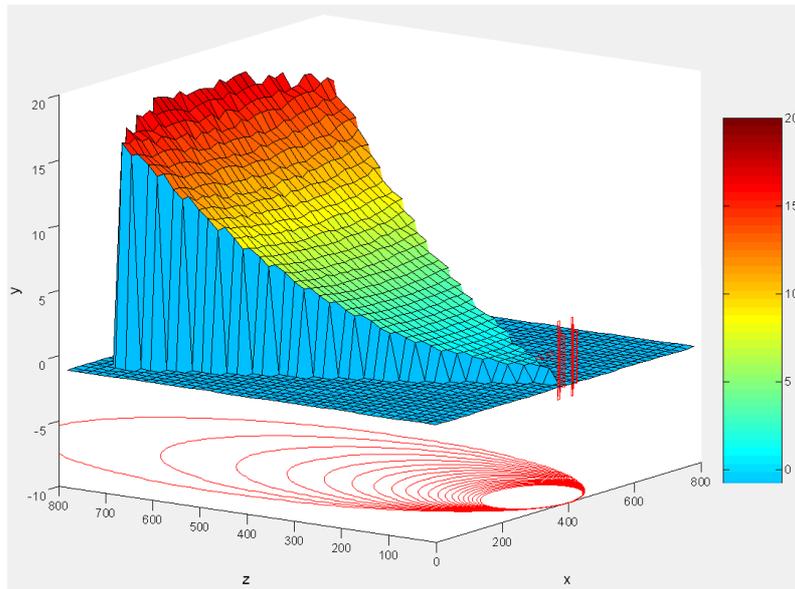

Fig. 4. The depth reconstruction uncertainty map for a stereo pair's FoV, where *B*=40 cm, *f*=3.5 cm, *p*=0.004 cm.
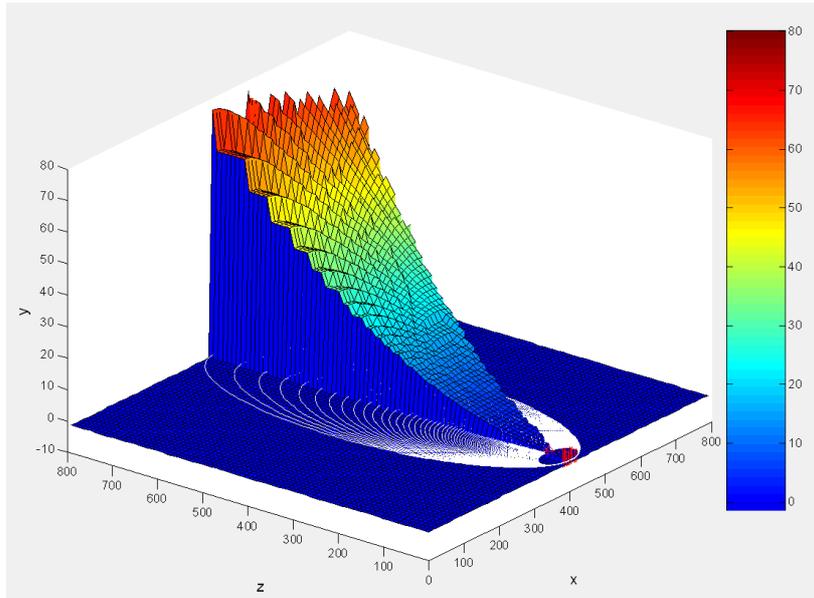
Fig. 5. The depth reconstruction uncertainty map for a stereo pair's half FoV, where B=20 cm, f=3.5 cm, p=0.008 cm.

Fig. 8(a) and Fig. 8(b) illustrate the variation of the uncertainty when both the focal length and the baseline length are changed for two different target distances, 200 cm and 600 cm, respectively. The uncertainty increases significantly when the baseline length decreases below 40 cm, independent of the location of the target within the FoV. Also, a significant increase in the uncertainty is visible for a focal length below 3.5 cm.

The relative accuracy is similar for a target located in different positions, but its absolute value is more significant for a target further from the stereo pair. In order to fulfil the reconstruction accuracy requirement for a faraway target, the focal length or baseline has to be adjusted. A longer focal length can be used to compensate for a shorter baseline. And also, in general, the longer the baseline is, the more difficult the matching becomes.

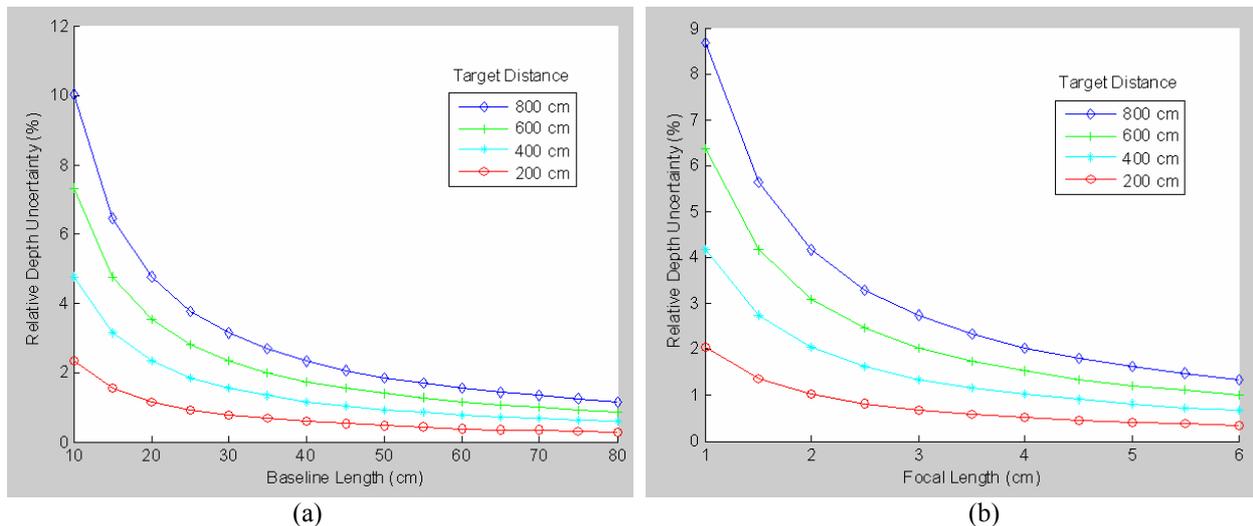(a)   (b)

Fig. 6. The uncertainty varies with the baseline length, focal length, sensor pixel length, and stereo convergence angle. The distance from the target to the camera is 800 cm, 600 cm, 400 cm and 200 cm, respectively. They are marked by different type of lines. The uncertainty varies with (a) the baseline length; (b) the focal length.

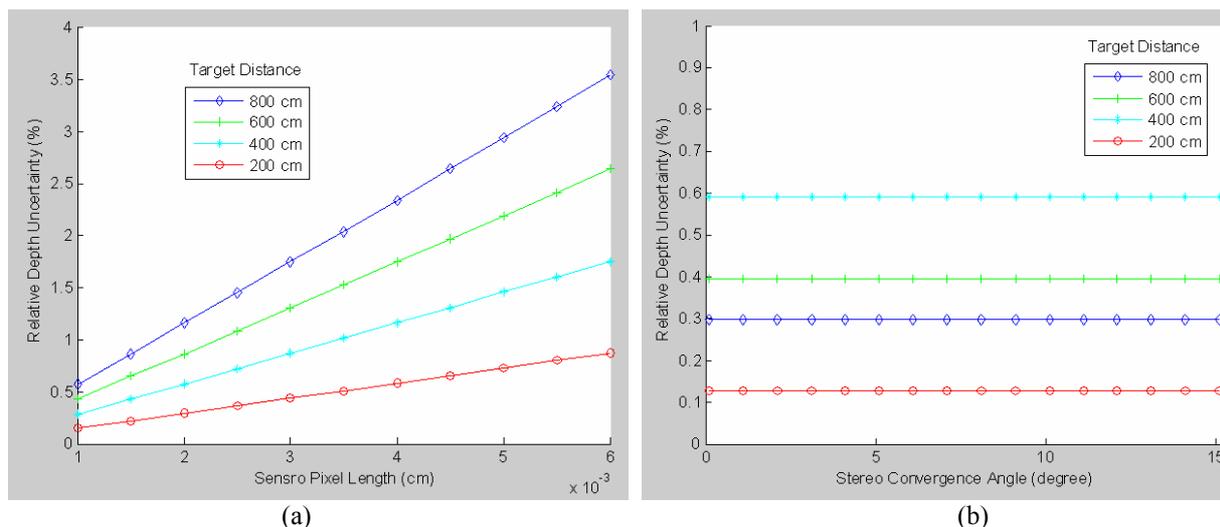(a)   (b)

Fig. 7. The uncertainty varies with the baseline length, focal length, sensor pixel length, and stereo convergence angle. The distance from the target to the camera is 800 cm, 600 cm, 400 cm and 200 cm, respectively. They are marked by different type of lines. The uncertainty varies with (a) the sensor pixel sizes; (b) the convergence angle.

6. CONCLUSION

The planning and control of a multi stereo pair's baselines, positions and poses for surveillance and tracking purposes, e.g. in supermarkets, museums, the home environment, and especially in situations which require stereo data to reconstruct 3D with a required accuracy, are possible fields of application. The proposed approach may be used in the dynamic control of stereo pair's baseline, and cameras' corresponding positions and poses, to observe a moving target.

The analysis presented shows that the depth reconstruction accuracy varies more significantly with respect to the target distance to baseline, baseline length and focal length than to the convergence angle. Small changes in stereo convergence angle do not affect the depth accuracy overly much, especially when the target is placed centrally. On the other hand it can have a great impact on the shape of the iso-disparity curves. From the proposed iso-disparity mathematical model we can get reliable control of the iso-disparity curves' shapes and intervals by using the systems configuration and target properties.

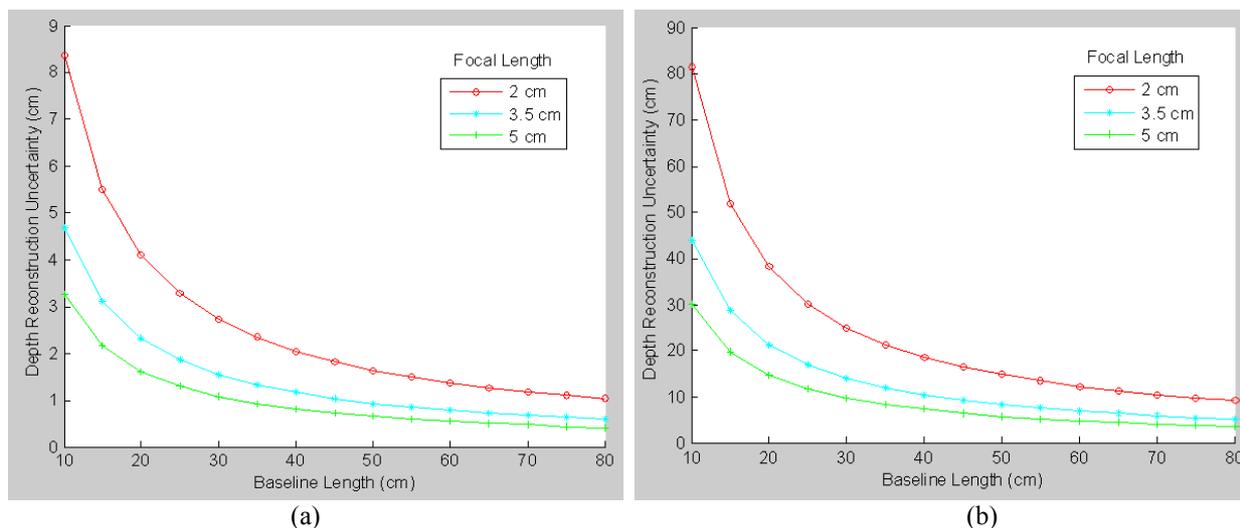(a)   (b)

Fig. 8. The uncertainty varies with both the focal length and the baseline length. The focal lengths are 2 cm, 3.5 cm and 5 cm, respectively, marked by different type of lines. The target is (a) 200 cm; (b) 600 cm faraway from the camera.

To achieve a more accurate 3D reconstruction of the target, it is better to bring the target to an area with a small depth reconstruction uncertainty. Furthermore, the controllable disparity distribution can determine and verify the assumptions which are used in stereo algorithms.

Future work could focus on the dynamical adjustment of the configuration of a stereo pair according to the target shape and position. The iso-disparity geometry model could also be used as a guide for stereo rectification or matching.

REFERENCES

[1]   M. Pollefeys and S. Sinha, "Iso-disparity Surfaces for General Stereo Configurations," in: Proc. of the 6th European Conf. on Computer Vision, 2004.

[2]   M. Christian, S. Robert, "Adaptive Area-based Stereo Matching," in: Proc. of SPIE, Three-Dimensional Image Capture and Applications, 3313, 14-24, 1998.

[3]   H. J. Kim, M. H. Yoo, and S.W. Lee, "Dynamic Vergence Using Disparity Flux," in: Proc. of the 1th IEEE Int. Workshop of the Biologically Motivated Computer Vision, 2000.

[4]   B. Völpel and W. M. Theimer, "Localization Uncertainty in Area-Based Stereo Algorithms," IEEE Transactions on Systems, Man, and Cybernetics, **25(12)** (1995).

[5]   J. Chen, S. Khatibi and W. Kulesza, "Planning of A Multi Stereo Visual Sensor System- Depth Accuracy and Variable Baseline Approach," in: Proc. of Int. Conf. of Capture, Transmission and Display of 3D Video, 2007.

[6]   Y. Nakabo, T. Mukai, Y. Hattori, Y. Takeuchi and N. Ohnishi, "Variable Baseline Stereo Tracking Vision System Using High-Speed Linear Slider," in: Proc. of the IEEE Int. Conf. On Robotics and Automation, 2005.

[7]   G. L. Mariottini, D. Prattichizzo, "The Epipolar Geometry Toolbox: Multiple View Geometry and Visual Servoing for Matlab," in: Proc. of IEEE Int. Conf. on Robotics and Automation, 2005.

[8]   K. Ogle, Researches in Binocular Vision, W.B. Saunders Company, Philadelphia & London, 1950.

[9]   J. Chen, S. Khatibi and W. Kulesza, "Planning of A Multi Stereo Visual Sensor System for A Human Activities Space," in: Proc. of the 2nd Int. Conf. on Computer Vision Theory and Applications, 2007.

[10] R. Halif and J. Flusser, "Numerically stable direct least squares fitting of ellipses," in: Proc. of the 6th Int. Conf. Computer Graphics and Visualization, 1998, pp. 125 – 132.