Iso-disparity surfaces for general stereo configurations

Marc Pollefeys and Sudipta Sinha

Department of Computer Science
University of North Carolina
Chapel Hill, N.C. 27599-3175
{marc,ssinha}@cs.unc.edu

Abstract. This paper discusses the iso-disparity surfaces for general stereo configurations. These are the surfaces that are observed at the same resolution along the epipolar lines in both images of a stereo pair. For stereo algorithms that include smoothness terms either implicitly through area-based correlation or explicitly by using penalty terms for neighboring pixels with dissimilar disparities these surfaces also represent the implicit hypothesis made during stereo matching. Although the shape of these surfaces is well known for the standard stereo case (i.e. fronto-parallel planes), surprisingly enough for two cameras in a general configuration to our knowledge their shape has not been studied. This is, however, very important since it represents the discretisation of stereo sampling in 3D space and represents absolute bounds on performance independent of later resampling. We prove that the intersections of these surfaces with an epipolar plane consists of a family of conics with three fixed points. There is an interesting relation to the human horopter and we show that for stereo the retinas act as if they were flat. Further we discuss the relevance of iso-disparity surfaces to image-pair rectification and active vision. In experiments we show how one can configure an active stereo head to align iso-disparity surfaces to scene structures of interest such as a vertical wall, allowing better and faster stereo results.

1 Introduction

In stereo matching pixels of one images are compared with pixels of another image to identify corresponding pixels. The accuracy at which this can be done is limited by the resolution of the images. Since the matching ambiguity is limited to the epipolar lines, it is the resolution along the epipolar lines that is relevant. Therefore, we argue that iso-disparity surfaces (where disparities are defined along the epipolar line) characterize the uncertainty and discretization in stereo reconstruction. While the geometry of these surfaces is very well known and understood in the standard stereo case (i.e. fronto-parallel planes located at distances inversely proportional to the disparity), only very little is known for general stereo configurations. However, as we will see later, the shape of these curves –and therefore what is achievable by stereo– depends dramatically on the geometric configuration (and the internal cameras settings). It is also important to be aware that image pair rectification can only affect negatively (or at best keep unchanged) the intrinsic uncertainty of the stereo reconstruction for a certain configuration.

In addition, many stereo algorithms make assumptions about the scene that are related to the iso-disparity surfaces. Although many different approaches have been proposed to compute dense stereo correspondences (we refer the readers to [22] for a recent

review), most of these algorithms share common properties. Many algorithms evaluate potential matches by comparing pixels within a support region located around the points of interest. This pixel-by-pixel comparison of the region makes an implicit assumption about the 3D geometry of the observed surface. In fact, in this case the stereo algorithm "sweeps" over the iso-disparity surfaces. These correspond to the shape hypothesis being considered by the stereo algorithm.

Besides this, many stereo algorithm perform an optimization over the epipolar line or over the whole image where not only the matching cost is minimized, but also the surface smoothness is taken into account. This allows to reduce problems with ambiguous matches. Often, a cost term of the form $\|\nabla d\|^2$ is used where d is the disparity (along the epipolar line) and ∇ represents the gradient operator. A similar term can also be used for computing dense disparity maps using optical flow [3]. This results in a bias towards 3D surfaces for which $\|\nabla d\| = 0$. In other words, we can conclude that in addition to characterize the uncertainty iso-disparity surfaces represent the implicit assumptions made by many stereo algorithms.

In Section 2 and 3 we will discuss the geometry of the iso-disparity surfaces. In Section 4 related work is discussed. In Section 5 the relevance to the human visual system is discussed and in Section 6 the impact on image-pair rectification is discussed. Section 7 discusses the application to active vision and the paper is concluded in Section 8.

2   The standard stereo case

Before analysing general stereo configurations, we will review the standard stereo case. The standard stereo configuration consists of two identical cameras with a relative displacement being a pure translation along the cameras' x-axes. In this configuration the the same scanlines from both images are corresponding epipolar lines. Therefore the stereo search can be limited to corresponding horizontal scanlines.

In this case the implicit stereo surfaces take on a particularly simple form. For a 3D point (X, Y, Z) and two cameras in the standard stereo configuration with focal length f and baseline b (and with the image plane aligned with the XY-plane), the observed disparity is

$$d = -\frac{fb}{Z} \ .$$

(1)

Therefore in this case the implicit stereo surfaces are fronto-parallel planes. This is a well-known result. An example is shown in Figure 1. The synthetic cameras are 20 cm apart, have parallel optical axis and a virtual image size and focal length of 1000 pixels (roughly equivalent to 35mm lense or 53 degrees field-of-view). To avoid clutter only one in ten iso-disparity curves is plotted and a dot is placed for one in ten pixels on those curves. The same camera and baseline are also used further in the paper.

3   The general case

Because the ambiguity caused by an unknown depth is limited to image displacements along the epipolar lines we define disparities in those terms. First we define a coordinate
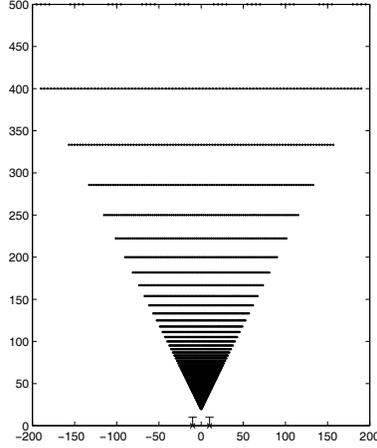
Fig. 1. Implicit stereo surfaces for the standard setup.

λ on the epipolar line where $\lambda(\mathtt{m}) = |\mathtt{m} - \mathtt{e}| - l_0$ where $\mathtt{e}$ is the epipole and with $l_0 = \min |\mathtt{m} - \mathtt{e}|$ over all image points $\mathtt{m}$. Note that λ is bound to $[0, \sqrt{w^2 + h^2}]$ (with w and h being the image width and height respectively). Now the disparity can be defined as $d = \lambda'(\mathtt{m}') - \xi\lambda(\mathtt{m})$ with $\xi \in \{-1, 1\}$ a term that takes into account a possible different orientation of the epipolar lines (see [13] for more details on this issue)[1]. A similar disparity measure was for example proposed in [3][2]. Note that this is different from the typical definition where disparities would be defined directly in difference of image coordinates, leading to both horizontal and vertical disparities, e.g. [23]. In the case where epipolar lines are horizontal both measures coincide.

Given the nature of the problem it is most relevant to start with the analysis of the shape of the iso-disparity curves wihtin epipolar planes. Let us call such a curve ϕ. Let us also define the point $\mathtt{M}_\infty$ as the point that is imaged at infinity in both images, i.e. the point that is on the intersection of both planes parallel with the image planes and passing through the centers of projection (and the considered epipolar plane). The following theorem can be proven:

Theorem 1. *The iso-disparity curves ϕ are conics that pass through both centers of projection and the point $\mathtt{M}_\infty$.*

Proof: Consider two centers of projection $\mathtt{C}_1$ and $\mathtt{C}_2$ and the intersections $\mathtt{I}_1$ and $\mathtt{I}_2$ of an epipolar plane with the two image planes (see Figure 2). Take four image points $\mathtt{m}_0, \mathtt{m}_1, \mathtt{m}_i, \mathtt{m}_j$ with coordinates 0, 1, i, j along the line and the point at infinity $\mathtt{m}_\infty$ on the epipolar line $\mathtt{I}_1$ and similarly $\mathtt{m}'_0, \mathtt{m}'_1, \mathtt{m}'_i, \mathtt{m}'_j$ and $\mathtt{m}'_\infty$ on the corresponding epipolar

---

[1] ξ should be chosen so that the visible part of the iso-disparity surface is observed from the same side by both cameras, e.g. $\xi = -1$ for verging cameras where the epipoles are on different sides of the image.

[2] Note that the definition proposed in [3] has the undesirable implictly assumption that corresponding epipolar lines have similar orientations.
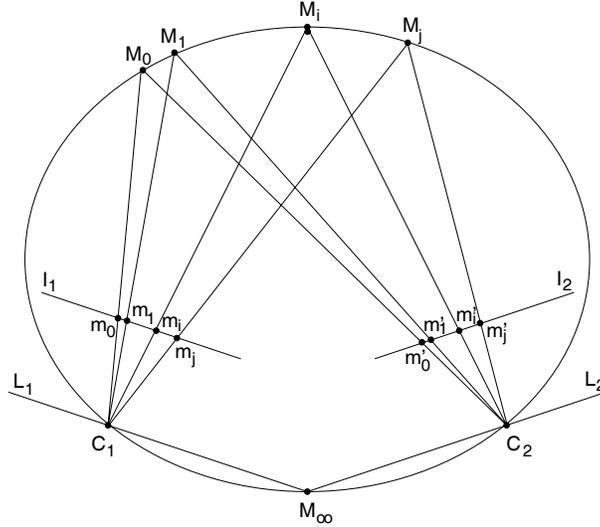
Fig. 2. Illustration of theorem

line I_2. Let us name the corresponding 3D points M_0, M_1, M_i, M_j and M_∞ respectively. Observe that the cross-ratio

$$\{M_0 C_1, M_1 C_1; M_i C_1, M_\infty C_1\} = \{m_0, m_1; m_i, m_\infty\}$$

is equal to

$$\{M_0 C_2, M_1 C_2; M_i C_2, M_\infty C_2\} = \{m'_0, m'_1; m'_i, m'_\infty\} \ .$$

Therefore, according to Chasles' theorem [7] both centers of projection C_1 and C_2 are on the same conic. The same analysis can be made when replacing M_i by M_j and since both conics share five common points M_0, M_1, M_∞, C_1 and C_2 they are equal. Note that since the cross-ratio is invariant to translation the proof is also valid for m'_0, m'_1, m'_i, m'_j having coordinates $d, d+1, d+i$ and $d+j$ respectively. $\qquad\square$

Note that with 3 fixed points, i.e. M_∞, C_1, C_2, there is a two degree family of possible conics left. One of the degrees of freedom corresponds to the actual disparity value that is considered, the other degree of freedom corresponds to the ratio of focal lengths for both images (i.e. the size of pixels).

It is also interesting to extend the analysis of the iso-disparity curves out of the epipolar plane. Let us for example consider all the image points with equal λ. By definition these form a circle with the epipole as center (and radius $l_0 + \lambda$). The corresponding points for a disparity d are located on a circle[3] with center e' and radius $l'_0 + \xi\lambda + d$. In general this type of iso-disparity curves will therefore be the intersection of two elliptic cones. Such a cone would be circular when the baseline is orthogonal to the image plane

---

[3] Depending on ξ it might be necessary to consider two separate half-circles. The separating line would be a line through the epipole and parallel with the other image plane. This is again related to the orientation of the epipolar lines [13].

and would become a plane when the baseline is parallel with the image plane. Note that –if we would desire to do so– the intersection of two cones is relatively simple to compute [14]. Since in our case the vertex of each cone is comprised within the other cone, a single connected intersection is obtained. If the epipoles are far from the images, the radii of the circles passing through the images are large and therefore the iso-disparity surfaces would have low curvatures (at least in the direction out of the epipolar planes). Some specific cases will be discussed further.

4 Related work

The special case for $d = 0$ is very much related with the concept of the *horopter* [1]. This curve is the locus of all 3D points that have identical image projections in both images of a stereo pair. Note that this is different from our definition of disparities since we are concerned with distance along the epipolar lines, but it can (partially) coincide in a number of cases. For the typical case of two cameras fixating a point, the horopter corresponds to an iso-disparity surface at least for the part contained within the horizontal plane. This part of the horopter is an ellipse similar to the one shown in Figure 2. In this case the vertical part of the horopter is a vertical line in the bisecting plane.

In fact, computer vision researchers have already used the concept of the horopter –or iso-disparity surfaces– in the context of stereo matching. Burt et al. [6] proposed to warp the images so that the horopter would better correspond to the expected scene content. By shifting scanlines the horopter of the standard stereo setup was transformed from the plane at infinity to the ground plane which is very useful in the context of robot navigation. We will show in Section 7 how this can also be achieved for a vertical wall by reconfiguring the camera. Others have also studied the possible use of the horopter in stereo vision, especially in the context of active stereo heads with verging [18] and torsion [12] degrees of freedom, or with asymmetrically skewed stereo cameras [8]. We will explore this possibility more in detail in the Section 7.

In [23] Völpel and Theimer also discuss iso-disparity loci and their relation to the reconstruction uncertainties in area-based stereo. However, instead of considering disparities along the epipolar lines, they consider independent disparities along the x- and y-direction. While our analysis is relevant to stereo algorithms that search for matches along epipolar lines, theirs is relevant to algorithms that look for corresponding pixels with similar coordinates without considering epipolar geometry. It is important to notice that this is fundamentally different and that both analyses only overlap when corresponding epipolar lines are aligned with corresponding image scanlines, i.e. for the standard stereo configuration and for a single scanline for some other stereo configurations. The analysis given in [23] is only meaningfull for stereo configurations that do not deviate to much from the standard stereo case.

5 Human Vision

The horopter has also been the subject of a lot of study in the area of human vision. Under the assumption that the retina is spherical and that corresponding points are

spread symmetrically around the fovea, the horopter becomes a circle known as the Vieth-Müller circle [17]. It has, however, been observed that the empirical horopter deviates from the theoretical horopter by what is known as the Hering-Hillebrand deviations [11]. The empirical horopter is always flatter than the Vieth-Müller circle. It is concave towards the viewer for close distances, becomes flat at a specific distance know as the abathical distance and then becomes convex. Ogle [16] proposed to model the empirical horopter using a conic section.

Using the concepts discussed earlier in this paper some interesting observations can be made. Since the horopter is flat at the abathical distance, the eye (at least along the eye's equator) can be modeled as a camera with a planar retina with the plane offset by an angle corresponding to the vergence required to fixate a point at the abathical distance. *In other words, the retinas acts as if they were flat!* As far as we could tell, this seems to be a new result [11]. Given a typical abathical distance of 1m and an interoccular distance of 6.5cm, the normal of the plane would deviate 1.86 degrees outwards from the visual axis. In this case Theorem 1 allows us to predict the shape of the horopter for any gaze direction and distance under the assumption that corresponding optical directions remain in correspondence (which is only an approximation given accomodation and other effects). If our eyes fixate a point straight in front of us, the 3 fixed points of Theorem 1, the fixation point and the need for symmetry are sufficient to determine uniquely the shape of the horizontal horopter. For excentric fixation, it can easily be shown that in the point of fixation the horopter should be tangent to the Vieth-Müller circle, i.e. the circle through both nodal points and the fixation point (given that the resolution of both fovea are the same), which in addition to the 4 known points also completely determines the shape of the horopter. This is illustrated with a few examples in Fig. 3.
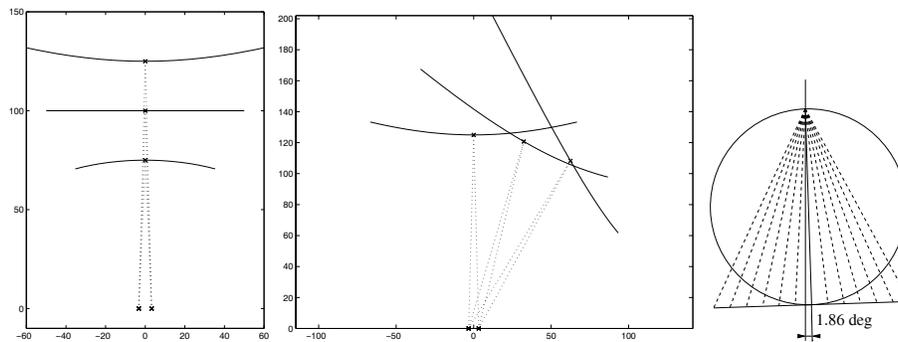
Fig. 3. Human horopter, as predicted by Theorem 1 (assuming an abathical distance of 1m and fixed corresponding points) for fixation at 75, 100 and 125cm (left) and for fixation at 125cm with 0, 15 and 30 degrees excentricity (middle). Simplified model of left eye with respect to stereo.

6 Rectification

To apply a stereo algorithm to a stereo pair that was not recorded using the standard stereo geometry, it is necessary to warp the images to make corresponding epipolar lines coincide with the scan-lines. This process is called *image-pair rectification*. In the past many different approaches have been proposed [4, 10, 15, 19, 9, 21, 20]. All these approaches somehow try to minimize distortions. Most of the approaches are homography based. For such approaches the epipole has to be warped to infinity to achieve parallel epipolar lines. In this case the resulting iso-disparity surfaces correspond to a set of planes, as in the standard stereo case or as the case with zoom discussed in Section 7.2. Clearly, when the original iso-disparity surfaces where far from planar, this type of rectification is bound to cause large distortions, at least for some parts of the images and of the iso-disparity surfaces.

Some recent approaches have abandoned the limitations imposed by homographies and have resorted to more general warping functions [21, 20]. Both approaches preserve distances along epipolar lines. Therefore, these approaches preserve the iso-disparity surfaces. A disadvantage is that these rectification approaches are computationally more expensive and do not preserve lines. Note, however, that nowadays arbitrary warping functions can efficiently be performed on graphics hardware using texturing functionality.

By using an iso-disparity preserving rectification and a stereo algorithm that computes integer disparity values, the reconstructed points will have to be located on a finite set of surfaces corresponding to the iso-disparity surfaces for integer disparities. Ideally, one would expect every scene point to be projected on the closest iso-disparity surface by the stereo reconstruction. In Figure 4 a point reconstruction is shown that was obtained using a rectification approach that preserves distances along epipolar lines. The reconstruction was obtained from the two castle images shown at the top of the figure using a stereo algorithm without subpixel matching. Note the clearly visible elliptic patterns in the top view of the reconstruction. Our second example consists of a forward moving camera. In this case it can be verified that the iso-disparity surfaces are predicted to be surfaces of revolution around the optical axis with parabola passing through both centers of projection as generators. This example is illustrated in Fig.5. The reconstruction was obtained from the two beguinage images shown on the left. The observed structures correspond to the prediction, i.e. the intersection of the 3D scene with the iso-disparity surfaces. While one can of course use an algorithm with subpixel precision, this doesn't change the fact that the depth precision of a stereo algorithm at a certain location will be proportional to the distance between the iso-disparity curves and that the algorithm will perform better if its priors (i.e. constant disparity) is aligned with the geometry of the observed scene.

7 Active vision

It has been established long ago [5, 2] that vision systems that would adapt themselves to the task at hand and/or to the observed environment can perform much better. The insights developed in this paper coupled with rectification algorithms that preserved the
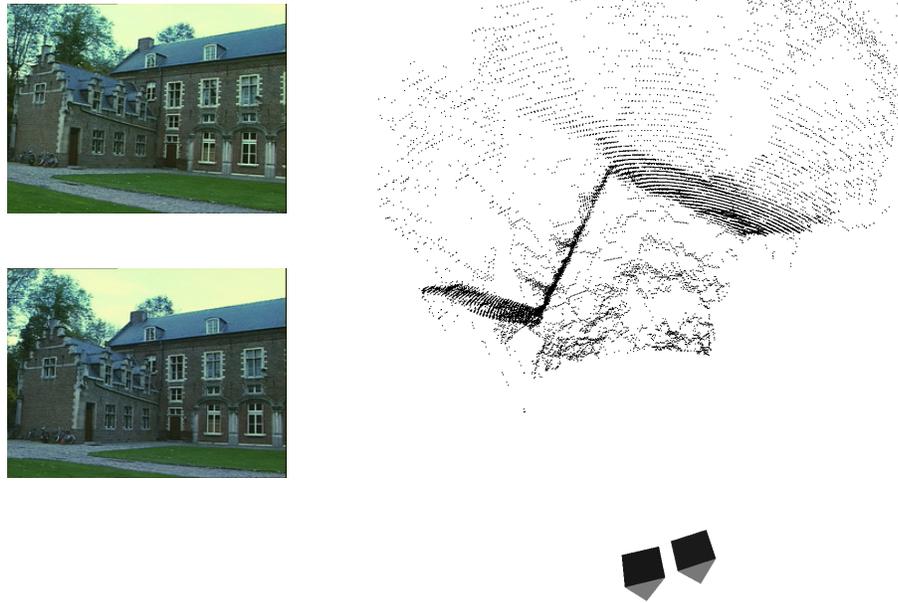
Fig. 4. Illustration of isodisparity surfaces for sideways motion (with some convergence). Image pair from the *castle* sequence (left) and 3D point reconstruction (right).

disparities can be very useful to provide the most optimal configuration for an active stereo head. Below we discuss the case of a verging stereo head and the case of a stereo head where the zoom of both cameras can be modified indepentently. Note that as mentioned earlier on todays hardware the expensive step for advanced rectification approaches is not the image warping anymore (since this can easily be achieved in real-time on graphics hardware), but the computation of the warping function. For an active vision scenario, one could compute the warping function once for many frames by keeping the configuration fixed for a number of consecutive frames, or, alternatively, precompute the warping function for a number of preset camera configurations. Those camera configurations would be matched to solve specific tasks, such as *reconstruct-left-wall, reconstruct-right-wall, reconstruct-object-in-front,* etc.

7.1 Vergence

Here we consider a typical active stereo head with verging capabilities. We study a synthetic configuration similar to the one shown in Figure 1, i.e we use the same virtual camera (width, height and focal length equal to 1000 pixels) and baseline (20cm). It is interesting to observe how small variations of the vergence angle cause important changes for the iso-disparity surfaces. In Figure 6 iso-disparity curves are plotted for the plane containing both optical axes with vergence going from 10 degrees divergent
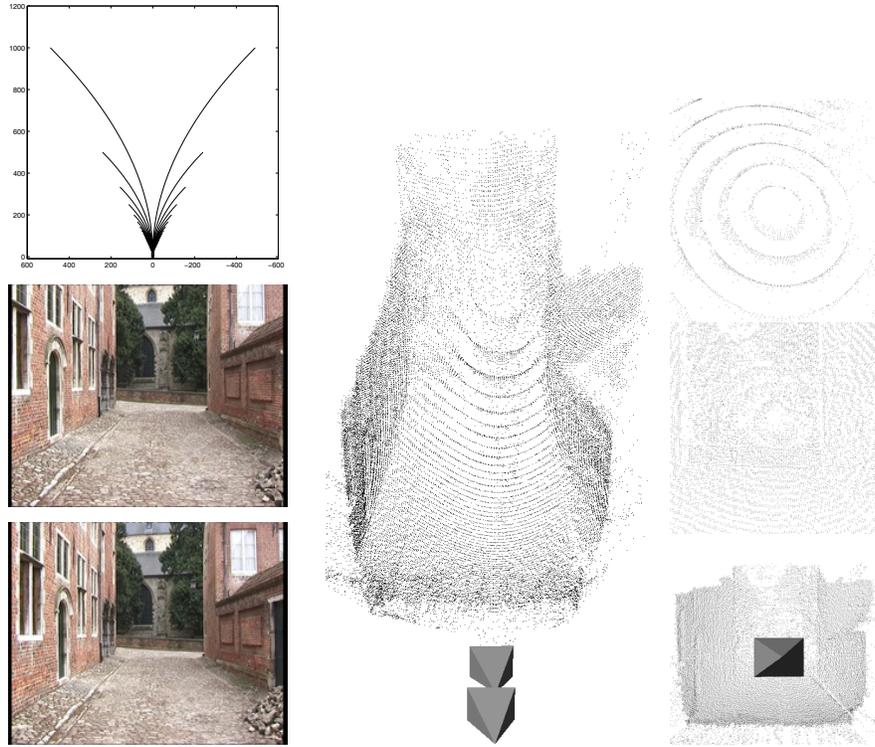
Fig. 5. Iso-disparity surfaces for forward motion. Top view of theoretical iso-disparity curves (upper-left). Notice that as expected reconstruction uncertainty is very large close to the direction of motion. Image pair from the *beguinage* sequence (lower-left) and different views of the 3D point reconstruction (middle/right). The images on the right are taken at different locations along the optical axis of the cameras.

to 15 degrees convergent. It seems that divergent setups are better suited for observing objects (as the iso-disparity contours would more closely follow the expected object surface), convergent setups are better for exploring an environment. Note for example how some iso-disparity curves would be nicely aligned with a 2m wide corridor in the 10° convergent case. A possible limitation for divergent configurations is the more limited overlap. However, as can be seen in this figure, for cameras with a field of view 53 degrees, a stereo field of view of 33 resp. 43 degrees is obtained for 5 and 10 degrees divergence. Note also that, by shifting the CCD in the camera similarly to the optical keystone found in many projectors, it would be possible to obtain a much larger overlap for divergent configurations.

7.2 Zooming

For active stereo zooming can also be used to change the shape of the iso-disparity surfaces depending on the task at hand. If the camera are kept parallel the iso-disparity
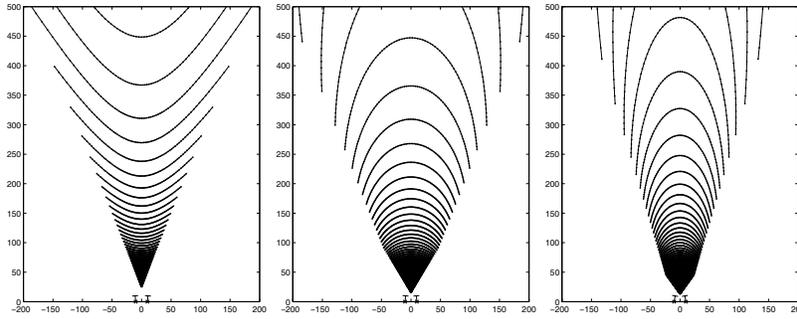
Fig. 6. Iso-disparity curves for different vergences, i.e. -5, 5 and 10 degrees for each camera (units are in cm).

surfaces remain a bundle of planes, but not fronto-parallel anymore. In Figure 7 a few examples are shown. Note how well zoomed in/out configurations are suited to reconstruct or follow a wall. In Figure 8 two stereo pairs are given recorded using a pair of
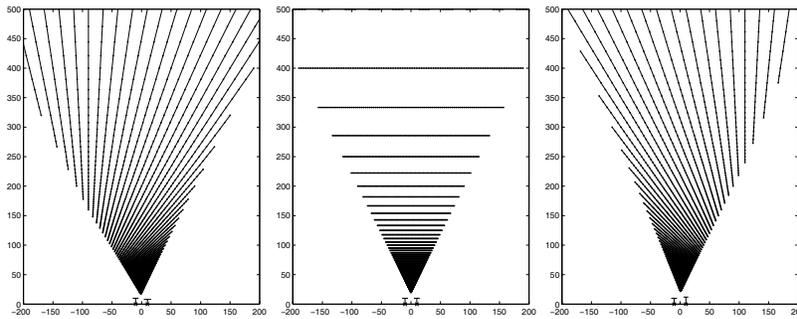
Fig. 7. Iso-disparity curves for different focal lengths (and parallel optical axis). The focal length of the left camera is 1000 and for the right camera 800, 1000 and 1200 (units are in cm).

pan-tilt-zoom cameras. The stereo pair is interlaced into a single image so that the disparities can be verified visually. The disparity images were computed using a standard stereo algorithm. The first stereo pair corresponds to the standard fronto-parallel stereo setup. The second stereo pair corresponds to the case shown on the left of Figure 7. One can clearly notice how for the second case the whole left wall fits within a very small disparity range. This benefits both the efficiency of the algorithm (less disparities to search) and the quality of the results (assumptions for area-based stereo are better met). Note that the same effect could be achieved by distording standard stereo images along epipolar lines, but in this case this would imply a loss of information or efficiency (to the contrary of the groundplane rectification case [6] where one could shift epipolar lines with respect to each other). More in general, this paper makes it possible for a robot equipped with a pair of pan-tilt-zoom cameras that explores an environment to

actively control its cameras to 'focus in' on areas of interest. This should be understood in the stereo sense, meaning bringing the structure of interest within a small disparity range. Note that small changes in intrinsics and/or orientation will not affect much the

Fig. 8. Stereo pair (interlaced to form single image) and corresponding disparity map for fronto-parallel case (top) and right-camera zoomed-out case (bottom, see Fig. 7, left).

depth discretisation of a particular stereo configuration. However, as we have shown it can have a great impact on the shape of the iso-disparity curves or in other words determine if the assumptions made by the stereo algorithm will be satisfied or not.

8   Summary and Conclusions

In this paper we have derived the shape of iso-disparity surfaces for general camera configuration. It was shown that the intersection of these implicit surfaces with epipolar planes have a particularly simple form. Comparing this with the empirical human horopter, we could conclude that the retinas act as if they were flat. This, as far as we were able to verify, is a new result and might lead to interesting further research in the area of human vision. The main goal of this paper was to provide more insight in the iso-disparity geometry for general camera configurations. We have discussed and illustrated the impact on rectification and advocate the use of disparity preserving rectification. Finally, we have shown how the insights developed in this paper can be exploited to 'focus' an active stereo head on a part of the scene. This allows to bring structures of interest within a small disparity range by controlling vergence and zooming, and thus allows to achieve better reconstruction of those structures faster.

Acknowledgments

References

1. F. Aguilonius, *Opticorum Libri Sex*, Ex off. Plantiniana apud viduam et filios Jo. Moreti, Antwerpen 1613.
2. Y. Aloimonos, I. Weiss, and A. Bandyopadhyay, "Active vision", *Proc. ICCV*, pp. 35-54, 1987.
3. L. Alvarez, R. Deriche, J.Sánchez and J.Weickert, "Dense Disparity Map Estimation Respecting Image Discontinuities: A PDE and Scale-Space Based Approach." *Journal of Visual Communication and Image Representation*. Vol. 13, No. 1/2, pp 3–21, March/June 2002.
4. N. Ayache, C. Hansen, "Rectification of images for binocular and trinocular stereovision", *Proc. International Conference on Pattern Recognition*, pp. 11–16, 1988.
5. R. Bajcsy, "Active Perception," *Proc. of the IEEE*, 76:996-1005, 1988.
6. P. Burt, L. Wixson, G. Salgian, "Electronically Directed 'Focal' Stereo", *Proc. ICCV*, pp. 97-101, 1995.
7. M. Chasles, *Traité des Sections Coniques.*, Gauthier-Villars, Paris, 1865.
8. A. Fransisco, F. Bergholm, "On the Importance of Being Asymmetric in Stereopsis–Or Why We Should Use Skewed Parallel Cameras", *IJCV* 29(3), 181-202 (1998).
9. J. Gluckman, S. Nayar, "Rectifying Transformations That Minimize Resampling Effects", *Proc. CVPR*, Vol. 1, pp. 111-117, 2001.
10. R. Hartley, "Theory and practice of projective rectification", *IJCV*, 35(2), pp. 115–127, November 1999.
11. I. Howard and B. Rogers, "Depth perception", *Seeing in Depth*, Vol. 2, Porteous, 2002.
12. M. Jenkin, J. Tsotsos, "Active stereo vision and cyclotorsion", *Proc. CVPR*, pp. 806—811, 1994.
13. S. Laveau and O. Faugeras. "Oriented projective geometry for computer vision". *Computer Vision - ECCV'96*, LNCS, Vol. 1064, Springer-Verlag, pp. 147-156, 1996.
14. J. Levin, "Mathematical models for determining the intersections of quadric surfaces", *Computer Graphics and Image Processing* 11(1):73-87, 1979.
15. C. Loop and Z. Zhang. "Computing Rectifying Homographies for Stereo Vision". *Proc. CVPR*, Vol.I, pages 125-131, 1999.
16. K. Ogle, "An analytic treatment of the longitudinal horopter; its mesaurements and the application to related phenomena especially to the relative size and shape of the ocular images", *Journal of Optical Society of America*, Vol. 22, pp. 665-728, 1932.
17. K. Ogle, *Researches in Binocular Vision*, W.B. Saunders company, Philadelphia & London, 1950.
18. T. Olson, "Stereopsis of Verging Systems", *Proc. CVPR*, pp. 55-60, 1993.
19. D. Papadimitriou and T. Dennis. "Epipolar line estimation and rectification for stereo images pairs", *IEEE Transactions on Image Processing*, 3(4):672-676, April 1996.
20. M. Pollefeys, R. Koch, L. Van Gool, "A simple and efficient rectification method for general motion", *Proc. ICCV*, pp. 496–501, 1999.
21. S. Roy, J. Meunier, I. Cox. "Cylindrical rectification to minimize epipolar distortion", *Proc. CVPR*, pp. 393–399, 1997.
22. D. Scharstein, R. Szeliski, "A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms", *IJCV*, Volume 47, Issue 1-3, pp. 7-42, April - June 2002.
23. B. Völpel, W.M. Theimer, "Localization Uncertainty In Area-Based Stereo Algorithms", T-SMC(25), 1995, pp. 1628-1634.