Fish-Eye-Stereo Calibration and Epipolar Rectification

Steffen Abraham [a], Wolfgang Förstner [b]

[a] *Robert Bosch GmbH, FV/SLH, PO-Pox 77 77 77, 31132 Hildesheim*

[b] *Universität Bonn, Institut für Photogrammetrie, Nußallee 15, 53115 Bonn, Germany*

Abstract

The paper describes calibration and epipolar rectification for stereo with fish-eye optics. While stereo processing of classical cameras is state of the art for many applications, stereo with fish-eye cameras have been much less discussed in literature. This paper discusses the geometric calibration and the epipolar rectification as prerequisite for stereo processing with fish-eyes. First, it surveys mathematical models to describe the projection. Then the paper presents a method of generating epipolar images which are suitable for stereo-processing with a field of view larger than 180 degrees in vertical and horizontal viewing directions. One example with 3D-point measuring from real fish-eye images demonstrates the feasibility of the calibration and rectification procedure.

Keywords: fish-eye camera calibration, fish-eye stereo, epipolar rectification

1 Introduction

Computer vision with a large field of view, with panoramic images and omnidirectional vision has received increasing attention (ECCV, 2002) especially for surveillance and navigation applications.

Different types of sensors are available for obtaining a large field of view: mirrors, rotating or moving cameras. A classical tool for acquiring images with a large field of view, common in surveillance applications, are fish-eye-optics. They provide images with a large field of view (1) with a single camera (2) from one view-point (3) at a single moment. Fish-eye-optics can be mounted on

Email addresses: Steffen.Abraham@de.bosch.com (Steffen Abraham), wf@ipb.uni-bonn.de (Wolfgang Förstner).

standard CCD- or CMOS-cameras without high technical effort. No external mirrors or rotating devices are required. Therefore the optics have a small dimension and do not need any maintenance.

Our paper deals with fish-eye-stereo. A stereo camera system equipped with two fish-eye optics can recover 3D-information in a large field of view around the cameras. These 3D-measurements are useful for obstacle detection in navigation or object recognition in surveillance applications.

The geometry of classical cameras can be well approximated by the pin-hole camera. This does not hold for cameras with a very large field of view. While stereo processing of classical cameras is state of the art for many applications, stereo processing with fish-eye cameras has been much less discussed in the literature.

Stereo processing usually requires two preparatory steps: (1) calibration of the stereo system including the determination of the interior and the relative orientation of the two cameras involved and (2) rectification of the image pair to obtain epipolar images, thus simplifying the correspondence problem. This paper discusses both problems in the context of fish-eye-stereo.

Geometric calibration of the fish-eye-stereo system requires a mathematical model to describe projection from 3D onto the 2D image plane.

One way is to extend the perspective projection equations for the pinhole camera with additional terms, e. g. polynomials, to compensate the fish-eye effect in the image plane. An example can be found in (Shah & Aggarwal, 1996). This type of model, however, is limited to a field of view much less than 180 degrees because the measurements in the image are corrected to a pinhole model with a planar image. Furthermore these models can lead to numerical instabilities in the calibration step due to inhomogeneous coverage of the virtual image plane.

A better way is to use a special fish-eye projection model consisting of two steps (1) an ideal map of the imaging sphere to a plane and (2) a subsequent correction.

In Section 2 we present a review of known fish-eye projection models. They can be derived from the optical design (Ray, 1994). The different models presented here have one projection centre (pinhole) in 3D, where all projection rays meet. Such models are proposed for calibration among others in (Fleck, 1995), (Xiong & Turkowski, 1997), (Bakstein & Pajdla, 2002). Our main goal is to integrate the fish-eye projection models into a framework for binocular vision with fish-eyes lenses.

In Section 3 we propose rectification models for fish-eye-stereo. Automatic stereo is simplified if epipolar lines of the two images are equal. Epipolar rectification can be seen as establishing a virtual camera system with ideal projection properties, namely the image planes being identical and parallel to the base line. With a perspective rectification the field of view is below 180 degrees, usually much lower, because of large distortions at the boundaries of the image.

Other proposals for rectification, found in the context of panoramic viewing e. g. (Ishiguro *et al.* , 1992) or omnivergent stereo e.g. (Shum *et al.* , 1999) refer to cylinders as projection surface. Binocular cylindrical panoramic images (Ishiguro *et al.* , 1992) limit the vertical field of view and do not lead to epipolar images. Multi-perspective panoramas or omnivergent stereo according (Shum *et al.* , 1999) do not have one projection centre for each camera and therefore cannot be used for our purpose. Our proposal can be seen as special case of the general rectification model given in (Pollefeys *et al.* , 1999).

Section 3 describes our concept for a virtual stereo fish-eye camera based on the fish-eye projection models from Section 2 and shows how epipolar stereo images for a large horizontal and vertical field of view can be obtained. The rectification models lead to omnidirectional binocular stereo images, which can be well adapted to the fish-eye projection geometry and which prevent too large image distortions during rectification. We give two examples in detail, one derived from the equi-distance projection model, the other from the stereographic projection model.

Section 4 presents a short survey of our calibration procedure. We show that calibration of a fish-eye stereo-system can be performed in a similar way as the well known calibration for perspective camera projections.

The system has been realized. The last Section 3.2.2 gives an example for deriving a point cloud from a fish-eye stereo system.

2 Mathematical models for fish-eye projection

2.1 Exterior and relative orientation

The first step of modelling fish-eye lenses is to establish a mapping from the viewing sphere to the image plane. The fish-eye projection with a field of view larger than 180 degrees requires non-linear transformations in addition to the projective or perspective geometry. As usual the projection of a 3D-point with world coordinates $\boldsymbol{X}_W = (X_W, Y_W, Z_W)^T$ into an image or sensor coordinate

$\boldsymbol{x}'_j = (x'_j, y'_j)$ is modelled as a sequence of coordinate transformations which contain the extrinsic and intrinsic camera parameters. We mainly follow the work of (Ray, 1994), (Fleck, 1995), (Xiong & Turkowski, 1997) and (Bakstein & Pajdla, 2002).

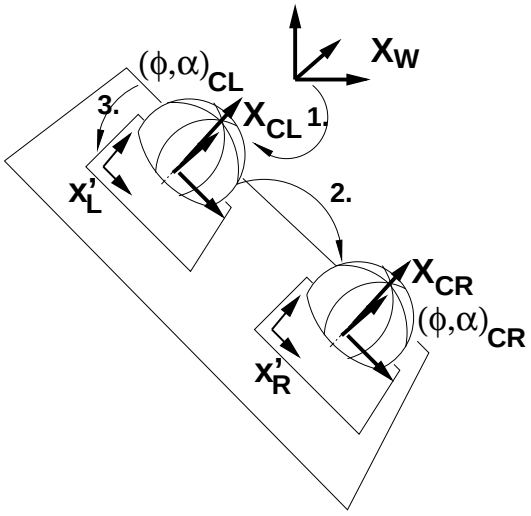

Fig. 1. Coordinate systems and subsequent transformations in the mathematical model for fish-eye stereo

The *exterior orientation* of the stereo system describes the coordinate transformation of the 3D-world coordinate $\boldsymbol{X}_W$ into the camera coordinate $\boldsymbol{X}_{CL}$ of the left camera using the rotation matrix $\boldsymbol{R}_{W,CL}$ and the translation vector $\boldsymbol{t}_{W,CL}$ as *extrinsic parameters* (cf. figure 1).

$$\boldsymbol{X}_{CL} \;=\; \boldsymbol{R}_{W,CL}(\boldsymbol{X}_W \;-\; \boldsymbol{t}_{W,CL}) \tag{1}$$

The *relative orientation* describes the coordinate transformation from the left to the right camera-coordinate system using the rotation $\boldsymbol{R}_{CL,CR}$ and the translation vector $\boldsymbol{t}_{CL,CR}$

$$\boldsymbol{X}_{CR} \;=\; \boldsymbol{R}_{CL,CR}(\boldsymbol{X}_{CL} \;-\; \boldsymbol{t}_{CL,CR}) \tag{2}$$

2.2 Interior Orientation

The transformation of 3D-camera coordinates $\boldsymbol{X}_C$ into image coordinates $\boldsymbol{x}'$ is separated into a *projection* and a *distortion model*. The projection model describes the ideal and error free projection.The distortion model describes deviations from the projection model, e. g. caused by lens distortions. Both models together contain the *intrinsic parameters*.

The fish-eye *projection models* in general are projections from a sphere onto a

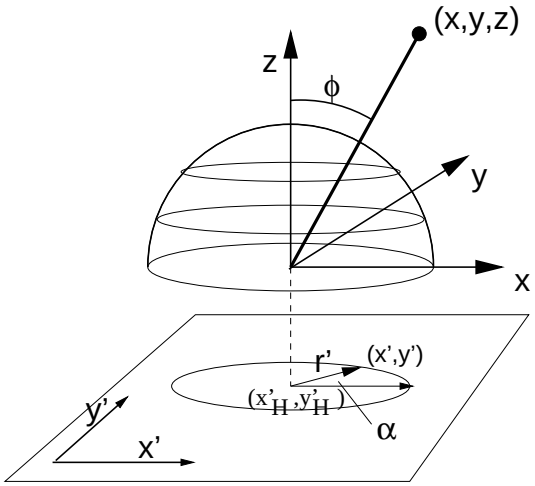

Fig. 2. Fish-eye projection, relation between coordinates (X, Y, Z) in the camera coordinate system, angles ϕ, α, radius r' and image coordinates (x', y')

plane, cf. Figure 2. The projection models are radial symmetric in relation to the optical axis. All models therefore refer to point (x'_H, y'_H) of best symmetry, being the intersection of the optical axis and the image plane. The distance r^* of an image point $\boldsymbol{x'}$ from the principal point then only depends on the angle ϕ between the optical ray from point to camera and the optical axis.

The angle α between the X_C-axis of the camera system and the projection of the ray onto the X_C and Y_C plane is transformed directly into the image. All intrinsic models $\boldsymbol{x'} = \boldsymbol{T(X)}$ therefore have the form:

$$\begin{aligned} x' &= c_x \cos[\alpha]\ r^*[\phi] + x'_H + \Delta x' \\ y' &= c_y \sin[\alpha]\ r^*[\phi] + y'_H + \Delta y' \end{aligned} \tag{3}$$

The radial projection function $r^*[\phi] = r^*[\arctan(\sqrt{X^2+Y^2}/Z)]$, the principal distances c_x and c_y [1] and the coordinates (x'_H, y'_H) of the principal point set up the projection model. The distortions from the projection model are encoded in the distortion model, $\Delta x'$ and $\Delta y'$, similar to perspective cameras. All parameters together form the *interior orientation* of the camera.

Normalizing the image coordinates in relation to the distortion parameters, the principal distance and the principal point coordinates give *normalized image* coordinates:

$$x^* = \frac{x' - x'_H - \Delta x'}{c_x}; \quad y^* = \frac{y' - y'_H - \Delta y'}{c_y} \tag{4}$$

[1] The principal distance is encoded here with 2 parameters c_x and c_y for a simpler description of video-cameras with non-square pixels

Table 1
Review of fish-eye projection models (cf. (Ray, 1994))

projection model	from camera to normalized image coordinates $(X,Y,Z) \rightarrow (x^*, y^*)$	inverse from normalized image to camera coordinates $(x^*, y^*) \rightarrow (X,Y,Z)$
perspective $r' = c\tan[\phi]$	$x^* = \frac{X}{Z}$ $y^* = \frac{Y}{Z}$	$X = x^*$ $Y = y^*$ $Z = 1$
stereo-graphic $r' = c\tan[\phi/2]$	$x^* = \frac{X}{\sqrt{X^2+Y^2+Z^2}+Z}$ $y^* = \frac{Y}{\sqrt{X^2+Y^2+Z^2}+Z}$	$X = \frac{2x^*}{1+x^{*2}+y^{*2}}$ $Y = \frac{2y^*}{1+x^{*2}+y^{*2}}$ $Z = \frac{1-(x^{*2}+y^{*2})}{1+x^{*2}+y^{*2}}$
equi-distant $r' = c\phi$	$x^* = \frac{X}{\sqrt{X^2+Y^2}} \arctan\left[\frac{\sqrt{X^2+Y^2}}{Z}\right]$ $y^* = \frac{Y}{\sqrt{X^2+Y^2}} \arctan\left[\frac{\sqrt{X^2+X^2}}{Z}\right]$	$X = \frac{x^*}{\sqrt{x^{*2}+y^{*2}}} \sin[\sqrt{x^{*2}+y^{*2}}]$ $Y = \frac{y^*}{\sqrt{x^{*2}+y^{*2}}} \sin[\sqrt{x^{*2}+y^{*2}}]$ $Z = \cos[\sqrt{x^{*2}+y^{*2}}]$
orthogonal $r' = c\sin[\phi]$	$x^* = \frac{X}{\sqrt{X^2+Y^2+Z^2}}$ $(Z>0)$ $y^* = \frac{Y}{\sqrt{X^2+Y^2+Z^2}}$	$X = x^*$ $Y = y^*$ $Z = \sqrt{1-(x^{*2}+y^{*2})}$
equi-solid-angle $r' = c\sin[\phi/2]$	$x^* = \frac{X}{\sqrt{2(X^2+Y^2)}}\sqrt{1-\frac{Z}{\sqrt{X^2+Y^2+Z^2}}}$ $y^* = \frac{Y}{\sqrt{2(X^2+Y^2)}}\sqrt{1-\frac{Z}{\sqrt{X^2+Y^2+Z^2}}}$	$X = 2x^*\sqrt{1-(x^{*2}+y^{*2})}$ $Y = 2y^*\sqrt{1-(x^{*2}+y^{*2})}$ $Z = \sqrt{1-(x^{*2}+y^{*2})}$

therefore $r^* = \sqrt{x^{*2}+y^{*2}}$. The entries in Table 1 refer to normalized image coordinates.

The standard perspective projection with $r^* = \tan[\phi] = \sqrt{X^2+Y^2}/Z$ is only a special case of the general model.

Table 1 collects the fish-eye projection models that can be found in the literature for technical optics (Ray, 1994), and in a few papers from computer vision (Fleck, 1995), (Xiong & Turkowski, 1997), (Bakstein & Pajdla, 2002). The table describes the transformations from camera into normalized image coordinates and the inverse transformations with an adequate normalization. To simplify the notation the 3D-camera-coordinates in Table 1 are used without suffix, thus $\boldsymbol{X}_C = \boldsymbol{X} = (X,Y,Z)^T$.

The *distortion model* $\Delta x'(x',y')$, $\Delta y'(x',y')$ can be set up with polynomials, describing radial symmetric, asymmetric or tangential distortion or can be

set up with orthogonal polynomials. Distortion polynomials are not subject of this paper. Proposals can be found in the literature e. g. (Fraser, 2001).

The mathematical models of this section are basis for the geometric calibration of a fish-eye camera system. For calibration one or combinations of these models (Bakstein & Pajdla, 2002) have to be selected depending on the characteristics of the camera. The determination of the parameters can be performed best with a self-calibrating bundle adjustment. Section 4 describes our procedure for calibration.

We are now prepared to discuss the epipolar rectification for fish-eye stereo.

3 Epipolar rectification for fish-eye stereo

3.1 Principle of epipolar rectification

Epipolar image rectification can be defined as geometric transformation of an image pair to an image pair which has the special property, that every scene point $\boldsymbol{X}_i$ is projected in both images into the same row ($y'_{Li} = y'_{Ri}$). i. e. such that the vertical parallax vanishes. This way the search for stereo correspondences can be reduced to a one-dimensional search along the image rows.

The image rectification process can be considered as a reprojection of the 3D-world into a *virtual stereo camera*. The virtual camera has ideal properties: parallel optical axes, identical interior orientation and no distortion. The projection centre of the real stereo and the virtual camera are kept at the same location. Figure 3 shows the geometry.

The concept of a virtual camera has some advantages. The rectified images are independent from the real projection system. The virtual camera can be designed for parallel epipolar lines. The image processing software does not need to know anything about the real camera and can be well adapted to the projection model of the virtual camera.

Following one of the models from Table 1 a virtual camera with the non-linear fish-eye projection onto the image plane, despite ideal properties like parallel optical axes, does not lead to epipolar images: epipolar lines are curves. Correspondence search in the images is complicated. Figures 4 and 5 show the projection of a simple 3D-object into a fish-eye stereo camera.

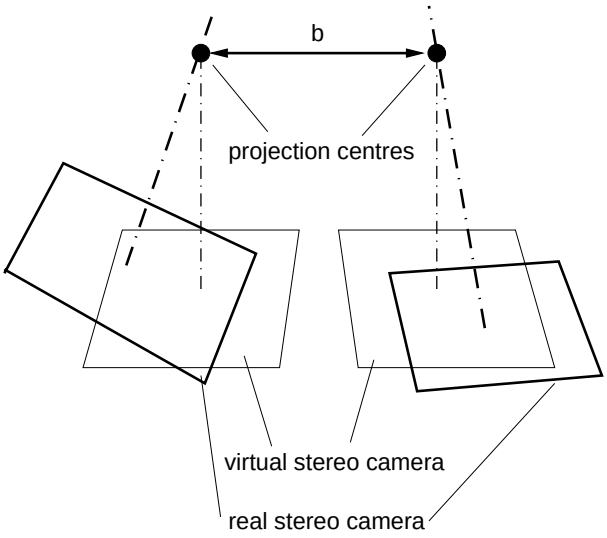

Fig. 3. Real and virtual stereo camera, rectification is identical to the projection of the world into the virtual stereo camera

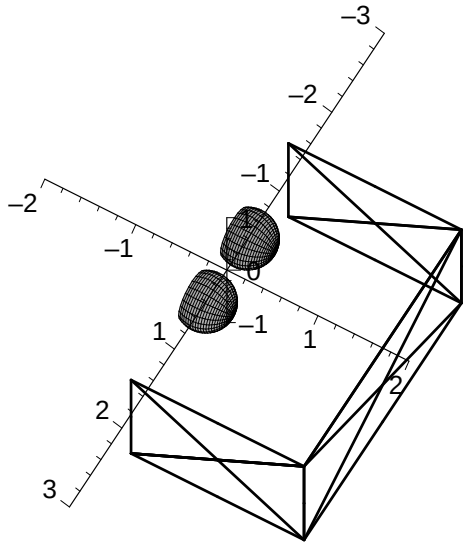

Fig. 4. Stereo camera system with fish-eyes, synthetic example, the fish-eyes are symbolized by the spheres looking into the rectangular object.

Going back to the basic properties of the epipolar geometry allows to define a rectification scheme useful for cameras with fish-eye lenses.

In a stereo camera system with two projection centres all epipolar planes form a pencil of planes in 3D-object space. The epipolar planes intersect in the baseline of the stereo system, cf. Figure 6. The orientations of the epipolar planes are characterized by the pitch angle β.

To get parallel epipolar lines in the rectified image pair every image row must correspond to an epipolar plane. To design a rectification model for fish-eyes one needs to setup an projection function $y^* = r_{y^*}(\beta)$ with $\beta = \arctan(Y_V/Z_V)$

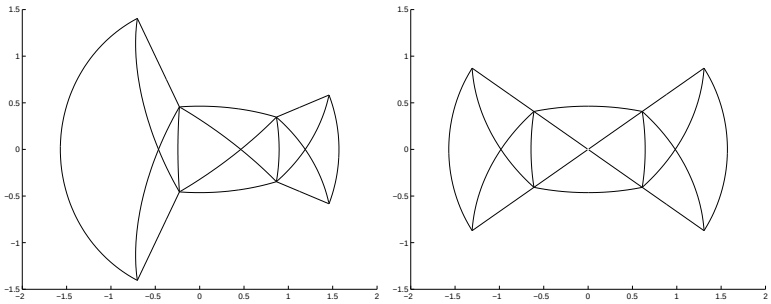

Fig. 5. Synthetic example from figure 4, projection into a fish-eye with the equi-distance projection model, left and right camera image with x' and y' image coordinate axis corresponding to β and ψ, unitless.

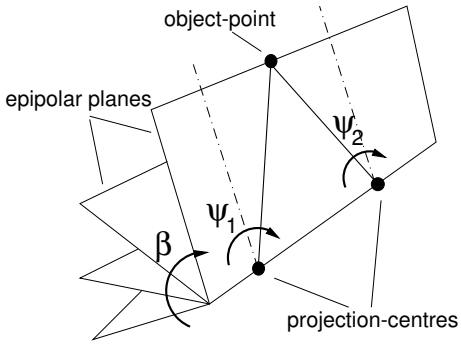

Fig. 6. Projection and epipolar planes

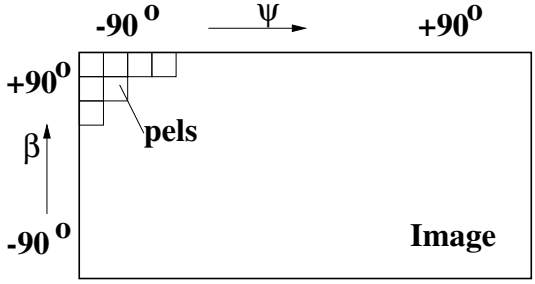

Fig. 7. The rectified image (β, ψ; cf. figure 6)

that defines the projection of the epipolar planes onto the rows of the rectified image. The projection function $x^* = r_{x^*}(\psi)$ with $\psi = \arctan(X_V/\sqrt{Y_V^2 + Z_V^2})$ defines the projection inside the epipolar planes, the projection inside the rows, cf. Figure 7. The two projection functions r_{x^*} and r_{y^*} for the virtual camera define the *rectification model* $\boldsymbol{x}'_V = \boldsymbol{T}_V(\boldsymbol{x}_V)$. The rectification model $\boldsymbol{T}_V$ is a non-linear function that transforms a coordinate $\boldsymbol{x}_V = (X_v, Y_v, Z_v)$ from the camera coordinate system into image coordinates $\boldsymbol{x}'_V = (x'_V, y'_V)$ of the virtual camera. The index V stands for the virtual camera. The rectification models are designed such that one obtains epipolar rectified images.

$$x'_V = c_{xV} \cdot r_{x^*}(\psi) + x_{HV}; \quad y'_V = c_{yV} \cdot r_{y^*}(\beta) + y_{HV} \tag{5}$$

Examples for projection models are given in Table 1. The functions r_{x^*} and r_{y^*} can be of different types. For example, the combination $r_{x^*} = \tan[\psi]$, $r_{y^*} = \beta$ can be found in (Roy *et al.* , 1997) as cylindrical rectification to minimize distortion effects. The projection function r_{x^*} can be identical or different for every image row. For practical reasons one would set the projection function r_{x^*} identically for all image rows.

The coordinate transformations for rectification can now be done according to the functional modelling in Section 2 by three transformation steps.

Rectification is a mapping of the real images into the virtual camera. Therefore for every pixel position of the virtual camera the gray value has to be taken from the image in the real camera. This means a transformation from virtual image coordinates into real image coordinates is required.

With the calibration model of the real camera (3) and the rectification model of the virtual camera (5) this coordinate transformation can be written in general as:

$$\boldsymbol{x}' \;=\; \boldsymbol{T}\left(R_{C,V}\;\boldsymbol{T}_V^{-1}\left(\boldsymbol{x}'_V\right)\right) \tag{6}$$

The rotation matrix $R_{C,V}$ defines the rotation between the camera coordinate systems of the real and the virtual camera (cf. Figure 3). The matrix $\boldsymbol{R}_{C,V}$ depends on the relative orientation of the stereo camera system (Fusiello *et al.* , 2000).

The concept of the virtual camera in conjunction with the different projection functions for the rectification process makes it possible to rectify images for stereo processing with a field of view larger than 180 degrees in vertical and horizontal directions. The rectification model can be selected according to the requirements. In most cases the projection model of the real camera and rectification model of the virtual camera should not differ too much in order to avoid large image deformations in the rectification process. This will be shown in two examples below.

3.2 Two rectification models for fish-eye-stereo

3.2.1 Epipolar equi-distance rectification model

First a modified equi-distance model with parallel epipolar lines is introduced. From the geometric relations depicted in the Figure 6 and the equi-distance-projection model in Table 1 the transformation of camera coordinates $\boldsymbol{X_C}$

into image coordinates $\boldsymbol{x}'$ can be derived:

$$\begin{aligned} x' &= c_x\psi \ + \ x'_H \qquad & y' &= c_y\beta \ + \ y'_H \\ \psi &= \arctan\left[\frac{X}{\sqrt{Y^2 \ + \ Z^2}}\right] & \beta &= \arctan\left[\frac{Y}{Z}\right] \end{aligned} \tag{7}$$

Image points with $y' = \text{const.}$ represent the epipolar plane where the object point can be found. The coordinate y' only depends on the angle β (cf. figure 6). The image coordinate x' depends on the location of the point in the epipolar plane. The angles β and ψ are projected in equidistant steps into the image $y' \sim c_y \cdot \beta$, $x' \sim c_x \cdot \psi$. The inverse transformation from an image coordinate

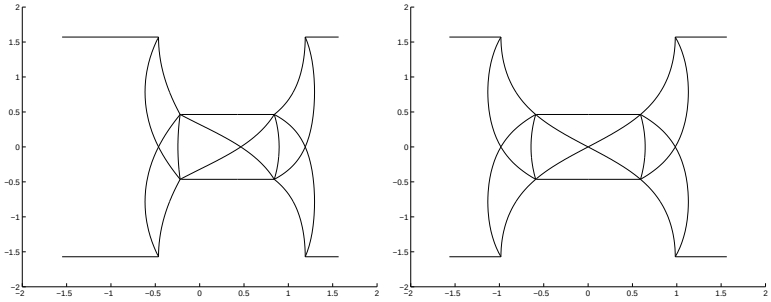

Fig. 8. Synthetic example from Figure 4, projection of the rectangular object with the derived *epipolar* equi-distance rectification model, left and right camera image with x' and y' image coordinate axis

$\boldsymbol{x}'$ into one camera coordinate $\boldsymbol{X}$ is given by:

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} \sin x^* \\ \cos x^* \sin y^* \\ \cos x^* \cos y^* \end{pmatrix} \tag{8}$$

with x^* and y^* as normalized image coordinates from (4). Figure 8 shows the rectification with the epipolar-equi-distant rectification model for the synthetic 'object' in Figure 4. An example of a real image pair can be found in Figure 11.

3.2.2 Epipolar stereographic rectification

As can be seen in Eqns. (7) and (8) the coordinate transformations always require the calculation of trigonometric functions (arctan(), sin(), cos()). For real-time processing this can be time consuming. Taking other fish-eye projection models from Table 1 into account trigonometric functions can be avoided.

For the *epipolar stereographic* rectification the set of transformation equations are from camera into image coordinates:

$$\begin{aligned} x' &= c \cdot \frac{X}{\sqrt{X^2+Y^2+Z^2}+\sqrt{Y^2+Z^2}} + x'_H \\ y' &= c \cdot \frac{Y}{\sqrt{Y^2+Z^2}+Z} + y'_H \end{aligned} \tag{9}$$

and from normalized image into camera coordinates:

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} \frac{2x^*}{1+x^{*2}} \\ \frac{1-x^{*2}}{1+x^{*2}} \frac{2y^{*2}}{1+y^{*2}} \\ \frac{1-x^{*2}}{1+x^{*2}} \frac{1-y^{*\prime 2}}{1+y^{*2}} \end{pmatrix} \tag{10}$$

Figure 11 shows an example applying the rectification model to a real image pair.

By rectifying images with the presented epipolar equi-distant or the epipolar stereographic rectification model, analogously to traditional epipolar rectification with the perspective model, a fast stereo correspondence search along lines can be performed. Except at the 'poles' where no stereo information is available a view of more than 180 degrees in horizontal and vertical direction can be processed. Fast rectification can be performed by pre-calculated look-up-tables.

4 Calibration

This section gives a short view inside our calibration procedure. We applied traditional algorithms from the field of photogrammetry and computer vision for calibration of fish-eye stereo. Our algorithm and most of the calibration algorithms generally consist of 3 steps: collecting images and performing automatic measurements, calculating approximate values with direct solutions and finally improving the start values with a non-linear optimization procedure.

Our calibration system performs self-calibration using a calibration board. Our calibration board here consists of three planes (not an requirement). The targeted points on the board are approximately known. The points are printed on paper by a laser printer, the papers have been attached to the planes. The board is imaged several times by the stereo-system and the imaged point patterns are measured. The parameters of the projection model and the 3D point coordinates are estimated simultaneously in a self-calibrating bundle adjustment.

Fig. 9. Example fish-eye-stereo: Original image pair and perspective epipolar rectification

The image processing software performs the point numbering and measurement automatically. Therefore, groups of four points are built twice on each plane allowing an automatic decoding (cf. figure 10). For each image approxi-

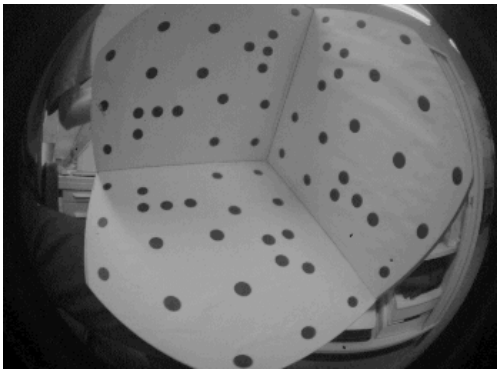

Fig. 10. Calibration target with point pattern

mate values of the exterior and interior orientation for the bundle adjustment are estimated by applying a modified Direct Linear Transformation algorithm that has integrated the equi-distant projection model. The approximately known 3D-target point coordinates are used as input in this calibration step.

After collecting measurements $\boldsymbol{x}'_{L_{ij}}$, $\boldsymbol{x}'_{R_{ij}}$ from all images j for the points i the intrinsic parameters $\widehat{\boldsymbol{p}}_{CL}$, $\widehat{\boldsymbol{p}}_{CR}$ of the left and the right camera, the relative or stereo parameters $(\widehat{\boldsymbol{R}}, \widehat{\boldsymbol{t}})_{CL,CR}$, the extrinsic parameters $(\widehat{\boldsymbol{R}}, \widehat{\boldsymbol{t}})_{W,CL_j}$ and the point coordinates $\widehat{\boldsymbol{X}}_{W_i}$ are estimated simultaneously in the non-linear

iterative self-calibrating bundle adjustment. The adjustment minimizes the reprojection error Ω between the projection model and the real measurements:

$$\Omega = \sum_i \sum_j \left[\boldsymbol{x}'_{L_{ij}} - \boldsymbol{f}\left(\widehat{\boldsymbol{p}}_{CL}, (\widehat{\boldsymbol{R}}, \widehat{\boldsymbol{t}})_{W,CL_j}, \widehat{\boldsymbol{X}}_{W_i} \right) \right]^2 + \left[\boldsymbol{x}'_{R_{ij}} - \boldsymbol{f}\left(\widehat{\boldsymbol{p}}_{CR}, (\widehat{\boldsymbol{R}}, \widehat{\boldsymbol{t}})_{W,CL_j}, (\widehat{\boldsymbol{R}}, \widehat{\boldsymbol{t}})_{CL,CR}, \widehat{\boldsymbol{X}}_{W_i} \right) \right]^2$$

In order to obtain best results, a selection of the fish-eye projection models from Table 1 is required. We select the one which fit the measurement data best based on the estimated mean reprojection errors $\widehat{\sigma}_{x'} = \sqrt{\Omega / r}$ with the redundancy r of the system.

Neglecting the fish-eye projection model the calibration procedure is equal to standard bundle adjustment self-calibrations. The calibration procedure can also be performed as test-field calibration with exactly known point field coordinates.

As final step we set up a rectification model and calculate Look-Up-Tables for image rectification.

5 Example

An example demonstrates the calibration, the rectification and the 3D-reconstruction for a stereo camera with fish-eye-lenses. Figures 9 and 11 show the original image pair and rectified images with different rectification models.

The calibration of the camera system was performed according Section 4. The radial projection function $r' = \sin[\phi/2]$ was selected for the stereo system. Table 2 gives the estimated calibration parameters with their standard deviations. The estimated mean reprojection error was $\widehat{\sigma}_{x'} = 0.13$ [pel].

Figure 9 shows the traditional perspective epipolar rectification on image planes. The image is obviously distorted strongly. With a larger field of view the performance of searching for stereo correspondences will be low.

Figure 11 shows epipolar rectifications with the epipolar equi-distance and the epipolar stereographic model. The rectified images are much less distorted than after the rectification with the perspective model shown in Figure 9. So for stereo-matching a much better performance could be reached than using the perspective rectification.

Comparing the epipolar equi distant and the stereographic rectification in Figure 11, the resulting images differ only slightly. In experiments both rec-

Table 2
Estimated calibration parameters with standard deviations, projection model $\sin[\phi/2]$

intrinsic parameters

	left	right
c_x $[pel]$	308.8 (0.5)	311.0 (0.6)
c_y $[pel]$	308.3(0.5)	310.7 (0.5)
x'_H $[pel]$	245.78 (0.03)	251.67 (0.02)
y'_H $[pel]$	129.20 (0.03)	125.56 (0.02)

relative orientation

t_x $[mm]$	78.64 (0.05)	r_x	0.17^o (0.01^o)
t_y $[mm]$	0.42 (0.03)	r_y	1.56^o (0.02^o)
t_z $[mm]$	0.62 (0.06)	r_z	0.15^o (0.01)

Fig. 11. Example fish-eye-stereo: Epipolar rectification with the epipolar-equi-distance ($r' = c \cdot \phi$) (top row) and the epipolar stereo-graphic $r' = c \cdot \tan[0.5 \cdot \phi]$ rectification model (bottom row)

tifications show nearly identical performance for stereo matching and both models are well suited for the investigated fish-eye optics.

The result of a 3D-reconstruction from the example image pair can be seen in Figure 12. The figure shows a horizontal 'slide' of the point-cloud. The slide is

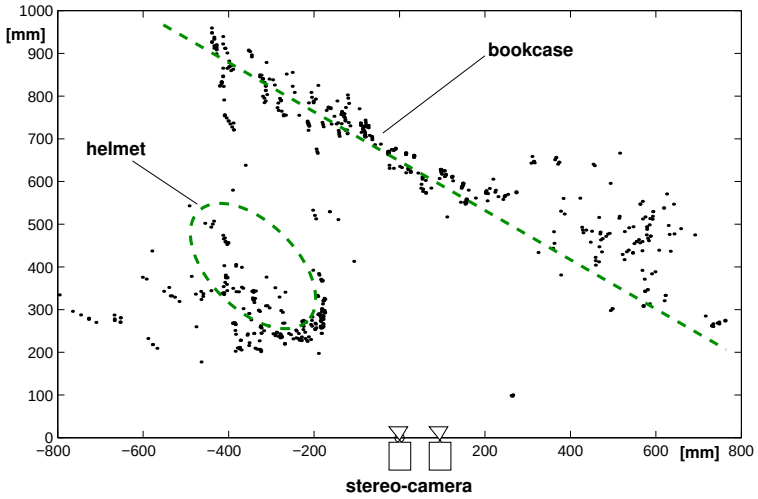

Fig. 12. Horizontal 'slide' of the 3D-point-cloud from fish-eye-stereo image pair in Figure 11,

indicated in Figure 11 bottom row with the white lines. In the point cloud the bookcase and the helmet can be clearly seen. The field of view is approximately 150 degrees. The field of view here is only limited by the optics not by the rectification model. The rectification model can also deal with fish-eye-optics with a larger field of view.

6 Conclusions

This paper has presented procedures for geometric calibration and epipolar rectification for stereo with fish-eye optics. Mathematical models for describing the projection by a fish-eye optics can be used to perform a geometric calibration of the stereo system. The rectification to epipolar stereo images has been considered as reprojection into a virtual camera. The virtual camera gives parallel epipolar lines, so a fast stereo processing with standard stereo matching algorithms is possible.

The rectification models for fish-eye images make it possible to handle a very large field of view with more than 180 degrees in vertical and horizontal directions from which 3D-measurements can be derived by classical matching algorithms. Since the projection properties of the rectification are similar to the real projection in the optics, the image deformations during rectification are small compared to traditional perspective rectification onto planes. This makes correspondence analysis with fish-eye lenses easier.

References

Bakstein, H., & Pajdla, T. 2002. Panoramic mosaicing with a 180^o field of view lens. *Pages 60–67 of: Third Workshop on Omnidirectional Vision*. IEEE Computer Society.

ECCV (ed). 2002. *Third Workshop on Omnidirectional Vision*. Held in conjunction with the ECCV: IEEE Computer Society.

Fleck, M. M. 1995. *Perspective projection: the wrong imaging model*. TR 95–01. Department of Computer Science, University of Iowa.

Fraser, C. S. 2001. Photogrammetric Camera Component Calibration: A Review of Analytic Techniques. *Pages 95–121 of:* Gruen, A., & Huang, T. (eds), *Calibration and Orientation of Cameras in Computer Vision*. Springer Series in Information Sciences, vol. 34. Springer, Berlin Heidelberg.

Fusiello, A., Trucco, E., & Verri, A. 2000. A compact algorithm for rectification of stereo pairs. *Machine Vision and Applications*, **12**(1), 16–22.

Ishiguro, H., Yamamato, M., & Tsuji, S. 1992. Omni-directional Stereo. *IEEE T-PAMI*, **14**(2), 257–262.

Pollefeys, M., Koch, R., & van Gool, L. 1999. A simple and efficient rectification method for general motion. *Pages 496–501 of: Proc. of ICCV*. IEEE Computer Society.

Ray, F. S. 1994. *Applied photographic optics*. Focal Press Oxford, UK.

Roy, S., Meunier, J., & Cox, I. J. 1997. Cylindrical rectification to minimize epipolar distortion. *Pages 393–399 of: Proc. of CVPR*. IEEE Computer Society.

Shah, S., & Aggarwal, J. K. 1996. Intrinsic parameter calibration procedure for a (high-distortion) fish-eye lens camera with distortion model and accuracy estimation. *Pattern Recognition*, **29**(11), 1–14.

Shum, H. Y., Kalai, A., & Seitz, S. M. 1999. Omnivergent Stereo. *Pages 22–29 of: Proc. of ICCV*. IEEE Computer Society.

Xiong, Y., & Turkowski, K. 1997. Creating image-based VR using a self-calibrating fis-heye lens. *Pages 237–243 of: Proc. of CVPR*. IEEE Computer Society.