The Inversion Camera Model *

Christian Perwass and Gerald Sommer

Institut für Informatik, CAU Kiel Christian-Albrechts-Platz 4, 24118 Kiel, Germany chp,gs@ks.informatik.uni-kiel.de

Abstract. In this paper a novel camera model, the *inversion camera model*, is introduced, which encompasses the standard pinhole camera model, an extension of the division model for lens distortion, and the model for catadioptric cameras with parabolic mirror. All these different camera types can be modeled by essentially varying two parameters. The feasibility of this camera model is presented in experiments where object pose, camera focal length and lens distortion are estimated simultaneously.

1 Introduction

In a typical application utilizing wide angle lens cameras, the cameras' images have to be rectified before they can be used. Various lens distortion models have been suggested for this purpose, like the widely used polynomial model [5], the bicubic model [7], the rational model [1] or the division model [3]. Another type of imaging systems that are particularly useful for navigation applications are catadioptric cameras, since they allow a 360 degree view in a single image. Geyer and Daniilidis showed in [4] how such systems can be modeled quite easily mathematically.

In this paper a novel camera model, the *inversion camera model*, is introduced, which combines the pinhole camera model, a lens distortion model and a model for catadioptric cameras with parabolic mirrors. As is shown later on, the lens distortion model is just the division model introduced by Fitzgibbon in [3] and the catadioptric camera model has been first presented by Geyer and Daniilidis in [4]. However, the authors found that both models can be represented in much the same way using inversion in a sphere. This also extends the division model to lenses with an angular field of view (FOV) of 180 degrees or more.

Inversion in a sphere can be represented as a (tri-)linear function in the Geometric Algebra of conformal space, which makes this algebra an ideal mathematical framework to work with the inversion camera model. The inversion camera model can be expressed as an algebraic entity of Geometric Algebra, i.e. a *multivector*, and a covariance matrix can be associated with it, which makes it directly applicable to statistical *linear* estimation methods as presented in [10, 9]. This is demonstrated in section 3, where results of the simultaneous estimation of object pose, camera focal length and lens distortion are presented.

^{*} This work has been supported by DFG grant SO-320/2-3.





(a) Basic pinhole camera setup.

(b) Pinhole camera setup with inversion.



(c) Projective camera with lens distortion.

(d) Parabolic catadioptric camera.

Fig. 1. Different cameras representable by inversion camera model.

A detailed understanding of Geometric Algebra is not necessary to follow the ideas presented in this paper. General introductions to Geometric Algebra can be found in [11, 6, 2]. Discussions of the application of Geometric Algebra to the estimation of geometric entities and operators, which are most closely related to this text, are [12, 10, 9].

The structure of this paper is as follows. First a general introduction to the inversion camera model is given, which is followed by a detailed discussion of the representation of lens distortion and parabolic mirror imaging systems. Finally, experiments on the simultaneous estimation of pose, focal length and lens distortion are presented to test the feasibility of the inversion camera model.

2 The Inversion Camera Model

The basic setup of the inversion camera model is shown schematically in figure 1 for the different imaging setups. Figure 1(a) shows the setup of the pinhole camera model. Point F is the focal point or optical center, point X is a world point and Y is the image of X on the image plane P_{img} . In a typical problem setup, the image point Y is given and the projection ray L has to be evaluated.

If the pinhole camera's internal calibration is given, the projection ray L can immediately be evaluated in the camera's coordinate frame.

In the inversion camera model this pinhole camera setup is represented as shown in figure 1(b). The sphere S_{inv} with center C_{inv} is used to perform an inversion of the image plane P_{img} which results in the sphere S_{img} . In particular, image point Y is mapped to point Z. In figure 1(b) the center C_{inv} of inversion sphere S_{inv} coincides with the focal point F. In this case the inversion of Y in S_{inv} results again in a point on the projection ray L, independent of the inversion sphere radius. Therefore, this setup is equivalent to the standard pinhole camera setup.

Figure 1(c) demonstrates what happens when the inversion sphere is moved below the focal point. Now the image point \boldsymbol{Y} is mapped to \boldsymbol{Z} under an inversion in \boldsymbol{S}_{inv} . The corresponding projection ray \boldsymbol{L} is constructed by \boldsymbol{F} and \boldsymbol{Z} and thus does not pass through \boldsymbol{Y} anymore. It will be shown later on that this results in a lens distortion model similar to the division model proposed by Fitzgibbon [3].

Simply by moving the inversion sphere S_{inv} and the image plane P_{img} , catadioptric cameras with a parabolic mirror can be modeled. This construction is shown in figure 1(d), and is based on work by Geyer and Daniilidis [4]. An inversion of image point Y in sphere S_{inv} generates point Z. In this case, it is equivalent to an inverse stereographic projection of Y on the image sphere S_{img} , which is how this mapping is described in [4]. The corresponding projection ray L is again the line through F and Z.

The image Y of a world point X generated in this way is equivalent to the image generated by a parabolic mirror whose focal point lies in F, as is shown in [4]. That is, a light ray emitted from point X that would pass through the focal point F of the parabolic mirror, is reflected down parallel to the central axis of the parabolic mirror. This is also indicated in figure 1(d). The reflected light ray intersects the image plane P_{img} exactly in the point Y.

While the construction for the parabolic mirror in terms of a stereographic projection has been known for some while, the authors recognized that the stereographic projection can be replaced by an inversion, which makes this model readily representable in the Geometric Algebra of conformal space (CGA). In the following the mathematical details of the inversion camera model will be discussed.

Mathematical Formulation In all calculations that follow, a right handed coordinate system is assumed, whereby e_1 points towards the right along the horizontal image plane direction, e_2 points upwards along the vertical image plane direction and e_3 points from the image plane center towards the focal point or optical center. This implies that objects that are in front of the camera will have a negative e_3 coordinate.

The geometric setup of the inversion camera model as presented in the previous section, can be modeled algebraically in CGA as follows. Like all transformations in Geometric Algebra, the image point transformation in the inversion camera model will be represented by a versor K. That is, if Y represents an image point, then Z := KYK is the transformed image point. As can be seen in figure 1 the point Z will in general not lie on the image plane. However, the goal is to find a K such that Z lies on the 'correct' projection ray. The transformed image point in the image plane can then be estimated by intersecting the projection ray with the image plane.

One of the simplest forms \boldsymbol{K} can take on is

$$\boldsymbol{K} = \boldsymbol{T}_s \, \boldsymbol{S} \, \boldsymbol{\widetilde{T}}_s \, \boldsymbol{D}, \tag{1}$$

where S is a sphere centered on the origin, T_s is a translator (translation operator) and D a dilator (isotropic scaling operator). This form was also found to behave well numerically. The dilator scales the image, which has the same effect as varying the focal length, if the inversion sphere $S_{inv} := T_s S \tilde{T}_s$ is centered on the focal point (cf. figure 1(b)). If the inversion sphere is not centered on the focal point, the dilator also influences the distortion. In the following, the transformation $K Y \tilde{K}$ is analyzed in some detail.

To simplify matters, it is assumed that T_s translates the inversion sphere only along the e_3 axis. Furthermore, S is a sphere of radius r centered on the origin. This is expressed in CGA as $S := e_o - \frac{1}{2}r^2 e_\infty$ and $T_s := 1 - \frac{1}{2}\tau_s e_3 e_\infty$. It may then be shown that $S_{inv} = T_s S \tilde{T}_s = s_1 e_3 + \frac{1}{2}s_2 e_\infty + e_o$, with $s_1 := \tau_s$ and $s_2 := \tau_s^2 - r^2$. The inversion sphere S_{inv} can thus be regarded as a vector with two free parameters, that influence the sphere's position along e_3 and its radius.

The dilation operator D for a scaling by a factor $d \in \mathbb{R}$ is given by $D = 1 + \frac{1-d}{1+d} E$, where $E := e_{\infty} \wedge e_o$. For brevity we define $\tau_d := -\frac{1-d}{1+d}$, such that $D = 1 - \tau_d E$. The image point transformation operator K is then given by

$$K = S_{inv} D = k_1 e_3 + k_2 e_{\infty} + k_3 e_o + k_4 e_3 E, \qquad (2)$$

with $k_1 := s_1, k_2 := \frac{1}{2} s_2 (1 - \tau_d), k_3 := 1 + \tau_d$ and $k_4 := -\tau_d s_1$.

In the model setup, the image plane P_{img} passes through the origin and is perpendicular to e_3 . That is, image points lie in the e_1-e_2 -plane. An image point will be denoted in Euclidean space by $\mathbf{y} \in \mathbb{R}^3$ and its embedding in conformal space by $\mathbf{Y} := \mathcal{C}(\mathbf{y}) \in \mathbb{G}_{4,1}$, that is $\mathbf{Y} = \mathbf{y} + \frac{1}{2} \mathbf{y}^2 \mathbf{e}_{\infty} + \mathbf{e}_o$. The embedded image point \mathbf{Y} is then mapped to the point \mathbf{Z} on the image sphere \mathbf{S}_{img} , via $\mathbf{Z} = \mathbf{K} \mathbf{Y} \mathbf{\widetilde{K}}$. Intersecting the line through the focal point \mathbf{F} and the transformed point \mathbf{Z} with the image plane gives the respective undistorted Euclidean image point $\mathbf{y}_d \in \mathbb{R}^3$. From a straight forward, if tedious calculation, it follows that

$$\boldsymbol{y}_{d} = \frac{-(s_{1}^{2} - s_{2}) d}{s_{1} (s_{2} - s_{1}) + (s_{1} - 1) d^{2} \boldsymbol{y}^{2}} \boldsymbol{y} = \frac{\beta}{1 + \alpha \boldsymbol{y}^{2}} \boldsymbol{y},$$
(3)

where $\alpha := \frac{(s_1-1)d^2}{s_1(s_2-s_1)}$ and $\beta := \frac{-(s_1^2-s_2)d}{s_1(s_2-s_1)}$. Note that y_d/β is the division model as proposed by Fitzgibbon in [3].

Typically, lens distortion models are used to remove the distortion in an image independent of the focal length or angular field of view (FOV) of the imaging system that generated the image. This is usually done by either requiring that lines which appear curved in the image have to be straight, or by enforcing multi-view constraints given a number of images of the same scene. The rectified image can then be used for any other type of application. For this purpose and for lenses with a FOV of at most 180°, the inversion model is equivalent to the division model.

However, here the applicability of the inversion model as a *camera model* is investigated. That is to say, the lens distortion of a camera system is modeled directly in the context of a constraint equation. This is shown in section 3 in the context of monocular pose estimation.

Focal Length and Lens Distortion Relationship The distortion generated by the inversion model as given in equation (3), has the effect that focal length and distortion are not independent, since α and β are not independent. The factor β mainly represents an overall scaling of the image, while α mainly influences the distortion. The exact relationship will be discussed in the following.

First of all, note that the interrelation of α and β does not represent a drawback as compared to the division model, if the image rectification is done independently and previous to any other calculations, as pose estimation. However, if the inversion camera model is used directly in a constraint equation as in equation (5) in section 3, then not every level of distortion can be rectified for every focal length or field of view (FOV).

The relationship between the transformed image point y_d and the initial image point y is given by the factor $\omega := \frac{\beta}{1+\alpha y^2}$, such that $y_d = \omega y$. The factor ω is therefore a function of the squared radial distance y^2 of an image point from the image center. The locations of constant ω in an image thus form concentric circles about the image center. These circles will be called *iso-circles* in the following. An iso-circle of particular interest in the analysis is the one that touches the upper and lower borders of the image, i.e. its radius is equal to half the vertical extent of the image. This particular iso-circle will be called *vertical* iso-circle and its radius will be denoted by ρ_v .

The value of ω for image points on the vertical iso-circle is directly related to the vertical angular field of view (vFOV). Note that the relation to the focal length is more complex if lens distortion is present, since the focal length is now a function of y^2 . That is, the focal length depends on the position of an image point in the image. It is therefore more useful to define an *effective* focal length (EFL) as the focal length of the image points on the vertical iso-circle.

The value of ω for image points on the vertical iso-circle will by denoted by ω_v and is given by $\omega_v = \frac{\beta}{1+\alpha \rho_v^2}$. The Euclidean position vector \boldsymbol{f} of the focal point is in the following parameterized as $\boldsymbol{f} = \tau_f \boldsymbol{e}_3$. That is, if the image is neither scaled nor distorted, τ_f is the focal length. It may be shown that the EFL \boldsymbol{f}_e is related to ω_v by $\boldsymbol{f}_e = \tau_f/\omega_v$.

It is possible to vary the inversion sphere center τ_s and the image scaling d such that the diagonal angular field of view (dFOV), i.e. the image distortion,



Fig. 2. Vertical vs. diagonal field of view (FOV) for pinhole model (middle, green graph), maximum trapezoidal (top, red graph) and maximum cushion (bottom, blue graph) distortion. Inversion sphere radius is 0.5.

is varied, while f_e and thus the vFOV are kept constant. This relationship is shown in figure 2.

Here $\tau_f = 1$, r = 0.5 and the image plane size was assumed to be $23.7 \times 15.6mm$, which is the CCD-chip size of a D70 digital SLR camera. The middle, green line shows the relation between the vFOV and the *diagonal* angular field of view (dFOV) for a standard pinhole setup. The top, red line gives the relation for maximum trapezoidal distortion and the bottom, blue line for the maximum cushion distortion. It was found that the maximum dFOV does not depend on τ_f or r. The location of the kink in the minimum dFOV plot does depend on the combination of τ_f and r, though.

To check, whether the inversion camera model can model actual lenses, the vFOV and dFOV of two lenses were measured and plotted. The first was the zoom lens SIGMA DC 18-125mm, 1:3.5-5.6 D, set to 18mm. This lens lies in the achievable distortion range of the inversion camera model.

The second lens was the Nikkor AF Fisheye 10.5mm, 1:2.8 G ED. This lens is a corrected fisheye, whereby the image does not appear circular but fills the whole image. This is achieved by obtaining a 180° FOV only along the diagonal and compressing the image more along the vertical than the horizontal. As can be seen in figure 2 the 10.5mm lens cannot correctly be represented by the inversion camera model. This is due to the different type of projection of fisheye lenses, which cannot be modeled by the inversion camera model. In the pose estimation experiments presented in section 3, it turns out that the inversion camera model approximates the 10.5mm lens well enough, though, to achieve good pose estimation results.

It is important to note that the above analysis is only an indicator whether a lens may be representable in the inversion camera model, since the actual lens distortion will in general be a more complex function. However, it was already shown in [3] that the division model, which is equivalent to the inversion model in the case of lens distortion, is a sufficiently good approximation of lens distortion for many applications.

Catadioptric Camera with Parabolic Mirror With respect to figure 1(d), the generation of image point \boldsymbol{Y} from world point \boldsymbol{X} via reflection at a parabolic mirror, can be represented mathematically by projecting \boldsymbol{X} onto the sphere \boldsymbol{S}_{img} followed by an inversion in the sphere \boldsymbol{S}_{inv} . In contrast to the inversion model setup for lens distortion, the focal point \boldsymbol{F} lies on the image plane in this case.

The relation of the physical parabolic mirror with respect to the mathematical setup is indicated in figure 1(d). In a standard setup the sphere S_{img} has unit radius and is centered on the focal point F of the parabolic mirror. The corresponding parabolic mirror then has to pass through the intersection points of S_{img} with the image plane P_{img} . The inversion sphere S_{inv} has to be centered on C_{inv} and has to pass through the intersection points of S_{img} with P_{img} . This fixes the radius of S_{inv} to be $\sqrt{2}$.

If the location and radius of the inversion sphere S_{inv} is fixed, the only free parameter left in the inversion camera model from equation (1) is the dilation, i.e. scaling of the image.

It may be shown that the relation between the image scaling d and the focal length of the parabolic mirror μ is given by $d = 1/(2\mu)$. The image point transformation operator for such a parabolic mirror setup is therefore $\mathbf{K} = \mathbf{S}_{inv} \mathbf{D}$ whereby

$$S_{inv} = -e_3 - \frac{1}{2}e_\infty + e_o, \qquad D = 1 + \frac{2\mu - 1}{2\mu + 1}E.$$
 (4)

3 Experiments

The accuracy of the inversion camera model as compared to other lens distortion models, is the same as that of the division model introduced in [3]. A comparison of a number lens distortion models, including the division model can be found in [1], where it is shown that the division model with one free parameter has a rectification quality that is comparable to a fourth order radial polynomial approach with two free parameters.

To demonstrate the feasibility of the inversion camera model in the context of an application, monocular pose estimation experiments were carried out. In these experiments not only the pose of a known object from a single camera view was estimated but also the camera's focal length and lens distortion. In the case of a catadioptric imaging system with a parabolic mirror, the object's pose and the mirror's focal length were computed.

The monocular pose estimation treated here, assumes that a model of the object is known, whose pose in space (location and orientation) is to be estimated. This model is given as a set of feature points, and it is also assumed that the correspondences between object feature points and image points are known.

Monocular Pose Estimation The problem is therefore to evaluate the transformation operator (motor) M, such that a model point X comes to lie on the projection ray of a corresponding image point Y. If lens distortion is present or a catadioptric imaging system is used, the image point has to be transformed to a rectified point Z, via $Z = KY\widetilde{K}$, where K implements the inversion camera model. Hence, the transformed model point X has to lie on the projection ray through the focal point F and Z. This can be formulated in CGA as

$$((\boldsymbol{K}\boldsymbol{Y}\boldsymbol{K})\wedge\boldsymbol{F}\wedge\boldsymbol{e}_{\infty})\wedge(\boldsymbol{M}\boldsymbol{X}\boldsymbol{M})=0.$$
(5)

If K is known, then this is basically the same as the pose estimation constraint in [13]. In contrast to [13] both operators M and K are estimated here using the same concepts as in [10,9]. That means, equation (5) is written as a multilinear equation which is quadratic in the components of M and K. This equation is then linearized so that a Gauss-Markov model can be employed to estimate M and K iteratively. The Gauss-Markov estimation is started from a very rough, automatically computed heuristic estimate and can be refined using Gauss-Helmert estimation. The whole estimation process is thus *automatic* and no a priori knowledge about a starting pose is necessary.

Experimental Setup Note that the simultaneous estimation of object pose, focal length and lens distortion is only numerically stable if the object has a sufficiently large appearance in the image and its extension along the optical axis is at least the same as its extension parallel to the image plane. For the following experiments the model of a house was used which was approximately $20 \times 15 \times 15$ cm in size (L×H×W).

This house model was moved by a robot arm in front of a stationary camera. Since the robot movements have a positioning uncertainty of below 1 mm, these positions can be used as ground truth. Note that the model was not rotated since an exact calibration of the rotation center with respect to the model was not possible. The model was translated in an area of approximately 50cm parallel to the image plane and 35cm perpendicular to the image plane. The closest approach of the object to the camera was approximately 10cm.

Neither an internal nor an external calibration of the cameras was carried out before the pose estimation experiments. However, the CCD-chip size in millimeters and its resolution in pixels were known and it was assumed that the optical axis passes at a right angle through the center of the CCD-chip.

Two different cameras were used. A Nikon D70 digital SLR camera with a pixel size of $7.8 \times 7.8 \mu m$ and a resolution of 3008×2000 pixel was used to take pictures with three different lenses: the zoom lens SIGMA DC 18-125mm, 1:3.5-5.6 D, set to 18mm, the Nikkor AF Fisheye 10.5mm, 1:2.8 G ED, and the Sigma 8mm 1:4.0 EX DG Circular-Fisheye. The other camera was a LogLux camera-link camera with a resolution of 1280×1024 pixel and a pixel size of $6.7 \times 6.7 \mu m$, which was used with the parabolic mirror catadioptric imaging system.

Eight markers were attached to the visible corners of the house model. The correspondences between these model points and their apparent positions in the images were found manually. The constraint equation given in equation (5) was then used to estimate the object pose and camera model parameters for each of the images taken. For each camera-lens setup the house was moved to the same six positions. Because no external calibration of the cameras with the robot arm was available, the pose estimation accuracy is measured as the difference between the true and the estimated object translations.

In fact, the 15 difference vectors between all pairs of the 6 house positions were evaluated separately for the true and the estimated house positions. Then the rotation was found that best aligned the true and estimated difference vectors. This is basically an external calibration of the camera. Two quality measures were then calculated. First, the root mean squared (RMS) Euclidean distance between the true and estimated aligned difference vectors and second, the RMS of the ratio of the Euclidean distance between the difference vectors and the length of the true vector. That is, the latter is the RMS percentage error.

The algorithm was implemented in CLUScript, an interpreted programming language, and was executed with CLUCalc [8]. The software was run on a 1.6GHz Pentium M processor. An optimized implementation in C++ may be expected to increase the execution speed by a factor of 10.

Results The results of the experiments are shown in table 1. It may be surprising that the pose estimation is most accurate for the fisheye lenses, which were found not to be exactly representable by the inversion camera model. This is because the house model only appeared in part of the image, whose distortion can be modeled quite well locally. Furthermore, the camera could be placed closer to the objects with the fisheye lenses (8mm, 10.5mm), than with the 18mm lens. The larger error in the LogLux camera results are mainly due to the low effective resolution when using a parabolic mirror. A 360 degree view is in this case mapped to a circular band in the image.

Note again that these pose estimation results were achieved without a full camera calibration. Instead focal length and lens distortion were estimated simultaneously with the object pose.

4 Conclusions

In this paper a novel camera model is introduced, the *inversion camera model*. It combines in a single model the standard pinhole camera model, the division model of lens distortion and the model of parabolic mirror imaging systems. The inversion camera model is based on the inversion of image points in a sphere, which can be be expressed in a straight forward manner in Geometric Algebra as a multilinear operator. This also implies that the camera model operator can be treated just like any other transformation operator in Geometric Algebra, as for example, a Euclidean transformation. Thus linear statistical estimation methods as presented in [10, 9] can be applied.

The experimental results show that this camera model can be employed successfully in the simultaneous estimation of object pose, and camera model para-

Camera/Lens	RMS Err.	Rel. RMS Err.	Mean Iter.	Mean Time
D70 / 8mm	2.63mm	1.48%	5.17	0.41s
D70 / 10.5mm	2.37mm	1.50%	5.50	0.41s
D70 / 18mm	5.51mm	3.12%	5.17	0.42s
LogLux / Cata.	7.87mm	3.66%	8.83	0.70s

Table 1. Experimental results of pose estimation.

meters in a 'half' calibrated camera setup. Next to the model's good behaviour in an actual application, it is also another example of the unifying nature of Geometric Algebra.

References

- 1. David Claus and Andrew W. Fitzgibbon. A rational function lens distortion model for general cameras. In CVPR (1), pages 213–219, 2005.
- Leo Dorst. Honing geometric algebra for its use in the computer sciences. In G. Sommer, editor, *Geometric Computing with Clifford Algebra*, pages 127–151. Springer-Verlag, 2001.
- 3. A. W. Fitzgibbon. Simultaneous linear estimation of multiple view geometry and lens distortion. In CVPR (1), pages 125–132, 2001.
- C. Geyer and K. Daniilidis. Catadioptric projective geometry. International Journal of Computer Vision, (45):223–243, 2001.
- R. I. Hartley and A. Zissermann. Multiple View Geometry in Computer Vision. CUP, Cambridge, UK, 2 edition, 2003.
- 6. D. Hestenes and G. Sobczyk. Clifford Algebra to Geometric Calculus: A Unified Language for Mathematics and Physics. Reidel, Dordrecht, 1984.
- E. Kilpelä. Compensation of systematic errors of image and model coordinates. International Archives of Photogrammetry, XXIII(B9):407–427, 1980.
- 8. C. Perwass. CLUCalc. WWW, http://www.clucalc.info/, 2005.
- C. Perwass and W. Förstner. Uncertain Geometry with Circles, Spheres and Conics, volume 31 of Computational Imaging and Vision, pages 23–41. Springer-Verlag, 2006.
- C. Perwass, C. Gebken, and G. Sommer. Estimation of geometric entities and operators from uncertain data. In 27. Symposium f
 ür Mustererkennung, DAGM 2005, Wien, 29.8.-2.9.005, volume 3663 of LNCS, pages 459–467. Springer-Verlag, Berlin, Heidelberg, 2005.
- C. Perwass and D. Hildenbrand. Aspects of geometric algebra in Euclidean, projective and conformal space. Technical Report Number 0310, CAU Kiel, Institut für Informatik, September 2003.
- C. Perwass and G. Sommer. Numerical evaluation of versors with Clifford algebra. In Leo Dorst, Chris Doran, and Joan Lasenby, editors, *Applications of Geometric Algebra in Computer Science and Engineering*, pages 341–349. Birkhäuser, 2002.
- B. Rosenhahn and G. Sommer. Pose estimation in conformal geometric algebra, part II: Real-time pose estimation using extended feature concepts. *Journal of Mathematical Imaging and Vision*, 22:49–70, 2005.