A Unified Framework for Line Extraction in Dioptric and Catadioptric Cameras

J. Bermudez-Cameo, G. Lopez-Nicolas and J.J. Guerrero

Instituto de Investigación en Ingeniería de Aragón
Universidad de Zaragoza, Spain
Email: bermudez@unizar.es, gonlopez@unizar.es, jguerrer@unizar.es

Abstract. Many of the omnidirectional visual systems have revolution symmetry and, consequently, they can be described by the radially symmetric distortion model. Following this projection model, straight lines are projected on curves called line-images. In this paper we present a novel unified framework to deal with these line-images directly on the image which is valid for any central system. In order to validate this framework we have developed a method to extract line-images with a 2-points RANSAC, which makes use of the camera calibration. The proposed method also gives the adjacent regions of line-images which can be used for matching purposes. The line-images extractor has been implemented and tested with simulated and real images.

1 Introduction

Line-images have been extensively used in computer vision. When a projection system is central, the 3D line and the view-point (optical center) lies on the same plane $\Pi$ and the projection is described by a vector normal $n$ to this plane $\Pi$. In general, any point $X$ contained in plane $\Pi$ is projected on the line-image and satisfies a nice constraint like $n^T X = 0$. In perspective cameras this constraint is transformed to the image plane resulting a 2D line. When the projection system is not perspective the relationship is not linear and the projected line-image is a curve.

Many approaches, e.g [1, 2], solve the constraint for collinear points in the unitary sphere. The intersection of the plane $\Pi$ with the unitary sphere is a great circle which is related with the image using the projection model. Instead of working on the sphere the problem can be tackled directly on the image. This approach has been extensively used in catadioptric images. The catadioptric line projection is modelled by the proposal of Geyer et al. [3]. In this particular case line-images are conics [4, 5]. For the case of fisheyes, line-images have not been extensively used.

Most of conventional and non-conventional cameras have revolution symmetry. Even when this constraint is not perfectly satisfied differences with the model can be encapsulated in an additional linear transformation. Main advantage of radially symmetric distortion is that can be used to model many different devices including perspective cameras, fisheyes and catadioptric systems.
Fig. 1. Catadioptric sphere camera model: The 3D point $X$ is projected onto the sphere. Then this point is backprojected to a normalized plane through a virtual optic center located a distance $\xi$ from the effective viewpoint. This point $\bar{x}$ is transformed to the image plane using the collineation $H_c$. Fisheye camera models: The radius of the point on the image is distorted by a function $\hat{r} = h(\phi)$.

As the line projection on the raw image is defined by more than two points it contains information about the calibration and distortion model and can be used to correct the image distortion [6–8]. These works are not about the line-image as geometric form, however the line-image is implicitly contained in them. In spite of not being expressed explicitly, the workspace used in Tardif et al. [9] is very related with the space in which a line-image is represented in our proposal.

Once the line-projection model is performed one direct application is line-image extraction. When the direct line projection model is known a Hough transform approach could be used for line-detection. This is the approach used for line-extraction with catadioptric systems in [10–12]. In [2] a scheme of split and merge is proposed to extract line-images in catadioptric systems. This approach use the inverse point projection model to back-project the points to the unitary sphere where the robust fitting is done. In [13, 14] a line-extraction method for hypercatadioptric systems solving the equation of the conic on the normalized-plane is proposed.

In this paper, we present a framework for line-image extraction in central systems following radially symmetric distortion models. This unified framework is a generalisation of the method presented by Bermudez-Cameo et al. in [13]. This generalisation expands the results obtained for hypercatadioptric systems to other catadioptric systems and dioptric systems with revolution symmetry. Explicit analytic expressions have been obtained for paracatadioptric, equiangular-
fisheye, stereographic-fisheye and orthogonal-fisheye models. We show an expression for the homogeneous line-image equation which is coherent with radially symmetric distortion models. The image-space in which the line-image is represented is similar to the space used in [9] for self-calibration. Main difference with this work is that we focused on the line-image and that in our proposal the distortion function is analytically solved for each projection model instead of having an empirical solution. In general our proposal is analytically solved when the inverse point projection model exists. The line-image homogeneous equation defines an algebraic distance measured in pixels which approximates the distance from any point of the image to the projected curve. We define a new robust method to extract line-images using this expression valid for catadioptric systems and fisheyes. We show the behaviour of the line-images and compare this model with the extension of the catadioptric sphere model for fisheyes presented in [15]. The extraction method is used to obtain the adjacent regions of image segments.

The rest of the paper is organized as follows. In Section 2 we describe the catadioptric sphere model and the fisheyes projection model. In Section 3 we present a unified description to represent line-images in revolution symmetry systems. In Section 4 we show the line-extraction method. In Section 5 we test the line extraction method for simulated and real images. Finally we present the conclusions.

2 Projection Models for Central Systems with Revolution Symmetry

When a projection system conserves symmetry around an axis it could be described using cylindrical coordinates. If the system is central the projected rays lie on a common point called fixed viewpoint $O$. In this case, both constrains are well represented by the spherical coordinate system. Let $X$ be a 3D point in homogeneous coordinates $X = (X Y Z 1)^T$. This point is transformed to the reference system of the camera in which the origin is the fixed viewpoint $O$ of the system and the $Z$-axis is aligned with the axis of revolution. This transformation consists of a rotation $R$ and a translation $t$, therefore the projection matrix is $P = (R|t)$.

The point is projected onto a unitary sphere around the viewpoint $O$ of the system. It is defined with two angular coordinates $\phi$ and $\varphi$ as,

$$x = (\sin \phi \cos \varphi, \sin \phi \sin \varphi, \cos \phi)^T. \quad (1)$$

Depending on the projection model this point is mapped on the image using different expressions. Notice that any point lying on the revolution axis is projected on a image point called principal point. If the camera is correctly aligned with the axis of revolution we can observe that the coordinate $\theta$ of a polar system in the image centred in the principal point, is related with the spherical coordinate $\varphi$ via the pixel aspect ratio $k_{par}$, as $\tan \theta = \pm k_{par} \tan \varphi$.

$^1$ The sign in this expression is used to model reflections in catadioptric systems
Catadioptric and dioptic systems are projection systems which conserve the revolution symmetry. Many projection models are used to model this devices. In the following descriptions we assume that image points are described in a reference centred in the principal point. We also assume that pixel aspect ratio is equal to one which is valid in digital imagery. A point in this reference system is denominated with the notation $\hat{x}$. The transformation from this reference to the final image coordinate system is the following,

$$
\begin{pmatrix}
u \\
v \\
1
\end{pmatrix} = 
\begin{pmatrix}
1 & s & u_0 \\
0 & k_{par} & v_0 \\
0 & 0 & 1
\end{pmatrix} \hat{x}.
$$

(2)

2.1 Projection Models for Catadioptric Systems

<table>
<thead>
<tr>
<th>Type</th>
<th>$\xi$</th>
<th>$\eta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parabolic</td>
<td>1</td>
<td>$2p$</td>
</tr>
<tr>
<td>Hyperbolic: $\cos\chi = \frac{d}{\sqrt{d^2 + 4p^2}}$</td>
<td>$\sin\chi = \frac{2p}{\sqrt{d^2 + 4p^2}}$</td>
<td></td>
</tr>
<tr>
<td>Planar:</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

$d$: distance between focus
$4p$: latus rectum

Fig. 2. Parameter of the unified sphere model for catadioptric systems.

Under the sphere camera model [3] all central catadioptric systems can be modelled by a projection to the unitary sphere followed by a perspective projection via a virtual viewpoint located a distance $\xi$ from the effective viewpoint (see Fig. 1). Let $\hat{\mathbf{x}} = (\hat{x}, \hat{y}, 1)^T$ be a point on an image referenced to the principal point and given the spherical coordinates $\phi$ and $\varphi$ of the corresponding point on the unitary sphere then,

$$
\hat{x} = \frac{f\eta \sin\phi \cos\varphi}{\cos\phi + \xi} \quad \text{and} \quad \hat{y} = -\frac{f\eta \sin\phi \sin\varphi}{\cos\phi + \xi}.
$$

(3)

In polar coordinates the point is described by $\hat{\theta} = -\varphi$ and

$$
\hat{r} = \frac{f\eta \sin\phi}{\cos\phi + \xi} = \frac{f\eta \tan\phi}{1 + \xi \tan^2\phi + 1}.
$$

(4)

The geometry of the projection system is described by parameters $\xi$ and $\eta$ which have a different definition depending on the system. In particular, when using hypercatadioptric systems the mirror parameters $\xi$ and $\eta$ are related via a single parameter $\chi$ which is related with the semi-latus rectum of the generational hyperbola and the distance between foci (see Fig. 2).
2.2 Fisheye Models

Several models are used to describe point projection in dioptric systems [16–18] depending on the manufacturing procedure of the lens. Assuming pixel square, these models are expressed in polar coordinates \( (\hat{r}, \hat{\theta}) \). For all these models \( \hat{\theta} = \varphi \) and the radius changes depending on the camera type (see Table 1).

Table 1. Fish-eye projection models.

<table>
<thead>
<tr>
<th>Model</th>
<th>( \hat{r} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Equiangular-Fisheye</td>
<td>( f \varphi )</td>
</tr>
<tr>
<td>Stereographic-Fisheye</td>
<td>( 2f \tan \left( \frac{\varphi}{2} \right) )</td>
</tr>
<tr>
<td>Orthogonal-Fisheye</td>
<td>( f \sin (\varphi) )</td>
</tr>
</tbody>
</table>

Some authors [15] have used the catadioptric sphere model to calibrate fisheye models. In the case of the stereographic projection both models are equivalent when \( \xi = 1 \) and \( \eta = 2 \). For other cases it is assumed that \( \xi > 1 \). As we will show in the following sections, the catadioptric sphere model and the rest of the fisheye models are not equivalent and it is only a good approximation when the field of view (FOV) is less than 180 degrees.

3 Unified Description for Line Projection in Central Systems with Revolution Symmetry

Let \( \mathbf{II} = (n_x, n_y, n_z, 0)^T \) be a plane defined by a 3D line and the viewpoint of the system \( \mathbf{O} \). The projected line associated to the 3D line can be represented by \( \mathbf{n} = (n_x, n_y, n_z)^T \). Then, the points \( \mathbf{X} \) lying in the 3D line are projected to points \( \mathbf{x} \). These points satisfy \( \mathbf{n}^T \mathbf{x} = 0 \). Using the spherical representation (1) and assuming that \( \theta = \pm \varphi \) (square pixel) this equality could be expressed as

\[
\sin \varphi (n_x \hat{x} \pm n_y \hat{y}) + n_z \hat{r} \cos \varphi = 0 .
\]  

(5)

With the change of variable \( \hat{\alpha} = \frac{n_x \hat{x} + n_y \hat{y}}{n_z} \) we can isolate the model parameters from the normal describing the line, obtaining the expression,

\[
\hat{\alpha} = -\hat{r} \cot \varphi .
\]  

(6)

Notice that \( \hat{\alpha} = \hat{\alpha} (\hat{r}) \), as a result of \( \varphi = h^{-1}(\hat{r}) \) when we have symmetry of revolution and square pixel. Therefore, the constraint for points on the line projection, in image coordinates for systems with symmetry of revolution is

\[
n_x \hat{x} \pm n_y \hat{y} - n_z \hat{\alpha} (\hat{r}) = 0 ,
\]

(7)

where \( \hat{\alpha} \) is a different expression for each camera model depending on the radius and the model parameters (see Table 2).
Table 2. $\alpha$ depending on the projection model

<table>
<thead>
<tr>
<th>Perspective</th>
<th>Para</th>
<th>Hyper</th>
<th>Equiangular</th>
<th>Stereographic</th>
<th>Orthogonal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Catadioptric</td>
<td>Catadioptric</td>
<td>Fisheye</td>
<td>Fisheye</td>
<td>Fisheye</td>
<td></td>
</tr>
<tr>
<td>$-f + \frac{c^2}{M P} - f p$</td>
<td>$-f + \cos \chi \sqrt{\frac{f^2 + f^2}{\sin \chi}}$</td>
<td>$-\hat{r} \cot \frac{c}{f}$</td>
<td>$\frac{c^2}{M T} - f$</td>
<td>$\sqrt{f^2 - f^2}$</td>
<td></td>
</tr>
</tbody>
</table>

3.1 Line-Image Curve Representation

Equation (7) is the homogeneous representation of the line projection on the image. There exist two particular cases common to all the projection models showed above. First we have the case in which 3D lines are coplanar to the revolution axis. In this case $n_z = 0$ and the resulting line-image is a radial straight line passing through the principal point.

$$n_x \hat{x} \pm n_y \hat{y} = 0.$$  \hspace{1cm} (8)

The second particular case happens when $n = (0, 0, 1)^T$. In this case the line-image is the projection of the vanishing line. This projection is a circle centred at principal point and with radius $\hat{r}_{VL}$. This radius depends on the calibration and differs with the projection model (see Table 3). The line-image equation in this case has the form,

$$\hat{\alpha} (\hat{r}) = 0.$$  \hspace{1cm} (9)

Table 3. $\hat{r}_{VL}$ for different projection models.

<table>
<thead>
<tr>
<th>Perspective</th>
<th>Para</th>
<th>Hyper</th>
<th>Equiangular</th>
<th>Stereographic</th>
<th>Orthogonal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Catadioptric</td>
<td>Catadioptric</td>
<td>Fisheye</td>
<td>Fisheye</td>
<td>Fisheye</td>
<td></td>
</tr>
<tr>
<td>$\infty$</td>
<td>$2 f p$</td>
<td>$f \tan \chi$</td>
<td>$f \frac{c}{2}$</td>
<td>$2 f$</td>
<td>$f$</td>
</tr>
</tbody>
</table>

The general form for a line-image is a curve. The catadioptric case has been deeply studied in [5], and it has been proven that the line-image is a conic. The stereographic case could be expressed directly in terms of a catadioptric projection. The ortographic line-image is also a conic but not in terms of a catadioptric projection. For the other cases in general the curve is not a conic.

In Fig. 3, we show a parametric representation of line-images depending on the elevation angle of the normal $n$ describing the projection plane of a 3D line. Each image has been simulated for a different device but with the same $\hat{r}_{VL}$. In all the cases line-images are well approximated by conics when the points of the segment are inside the limits of the vanishing line projection (FOV lower
Fig. 3. Representation of line-images depending on the projection model and the elevation of the normal $n$.

Fig. 4. Comparison of line-images using the equiangular projection model and the catadioptric sphere model with $\xi > 1$. 
than 180 degrees)[19]. However in the case of equiangular and orthogonal fisheye (b)(d) the line-images are not well fitted by conics when we are in the parts of the image corresponding to FOV greater than 180 degrees. In Fig. 4 we show a comparison of the line-images in an equiangular fisheye and an approximation using the catadioptric sphere model [15]. Notice that the line-image is well fitted inside the vanishing line projection but not outside. We also show how each pair of conics intersects in four points instead of two giving a sense of non-geometric coherence.

### 3.2 The Line-image Homogeneous Equation as a Measure of Distance

The homogeneous expression of the line-image (7) defines a family of curves located to an algebraic distance from the original curve.

\[
d(\hat{x}, \hat{y}) = n_x \hat{x} \pm n_y \hat{y} - n_z \hat{z}.
\]  

(10)

This algebraic distance is an approximation of the metric distance from a point to the line-image and is defined in pixels. When using an algebraic distance based on conics (i.e for hypercatadioptric systems \(d = \sqrt{x^T \Omega_{\text{hyper}} x}\)) is known that given a fixed threshold the region around the conic have a different thickness depending on the elevation angle of the vector \(n\). With our proposal the distance is a good approximation in regions close to the line-image.

In Fig. 5 (a) we show a comparison between the minimum distance of a point to the line-image (blue dotted) and the proposed algebraic distance (red) for hypercatadioptric images. The algebraic distance approximates the real distance in regions which are close to the line-image, therefore can be used to discriminate if a point lies on a line-image or not. In Fig. 5 (b) we show the same comparison but using the algebraic distance defined by the expression of a conic on the image \(d = \sqrt{x^T \Omega_{\text{hyper}} x}\). We can see how this distance does not approximate well the metric distance in regions close to the curve. We also show that this distance is lower than the metric distance in vertical lines but higher when the lines are horizontal. In practice that means that the thickness of a region defined by a threshold variates considerably if elevation of \(n\) changes.

Therefore we conclude that the proposed algebraic distance (10) is useful to discriminate if a point belongs to a line-image in catadioptric systems. However we have observed that in orthogonal systems the defined region is not constant. In this case it is necessary to use additional criteria to determine if a point lies on a conic or not (this will be detailed in Section 4.2).

### 4 Two Points RANSAC for Image Fitting

In this section we present a generalization of the method presented by Bermudez-Cameo et al. in [13] to fit line-images in central projection systems with revolution symmetry. First, we show how to define a line-image using two points and the calibration of the system. Then we describe the computation of the gradient used in RANSAC and how to robustly fit the line-image.
4.1 Line-Image Definition with Two Points

Having a collection of at least two points lying on a line-image we can build an homogeneous linear system using (7). The solution of this linear system is the normal \( n \) describing the projection plane of a 3D line. Depending on the device type, the way to compute the variable \( \hat{\alpha} \) differs (see Table 2).

\[
M \begin{bmatrix} n_x \\ n_y \\ n_z \end{bmatrix} = \begin{pmatrix} \hat{x}_1 \pm \hat{y}_1 \hat{\alpha}_1 \\ \hat{x}_2 \pm \hat{y}_2 \hat{\alpha}_2 \\ \vdots & \vdots & \vdots \\ \hat{x}_n \pm \hat{y}_n \hat{\alpha}_n \end{pmatrix} \begin{bmatrix} n_x \\ n_y \\ n_z \end{bmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.
\]  

(11)

The system is solved using a Singular Values Decomposition (SVD). In particular with two points and solving for \( n_x, n_y \) and \( n_z \) we have

\[
n_x = \hat{y}_1 \hat{\alpha}_2 - \hat{y}_2 \hat{\alpha}_1, \quad n_y = \pm (\hat{x}_2 \hat{\alpha}_1 - \hat{x}_1 \hat{\alpha}_2) \quad \text{and} \quad n_z = \hat{x}_2 \hat{y}_1 - \hat{x}_1 \hat{y}_2.
\]  

(12)

In contrast with [13], here points are defined in the image plane instead of the normalized-plane. Therefore, in our proposal the residual vector \( \delta = Mn \) is measured in pixel units.

---

\(^2\) The sign when ‘\( \pm \)’ in the following equations is positive for dioptic systems and negative for catadioptrics.

\(^3\) Points are referenced to the principal point. The normalized-plane is an intermediate projection plane described in the sphere-model (see Fig. 1).
4.2 Gradient of the Line-Image Curve

The gradient of the algebraic distance (10) is a vector perpendicular to the line-image in each point of the curve.

\[
\frac{\partial d}{\partial \hat{x}} = n_x - n_z \frac{\partial \hat{\alpha}}{\partial \hat{r}} \hat{r} \quad \frac{\partial d}{\partial \hat{y}} = \pm n_y - n_z \frac{\partial \hat{\alpha}}{\partial \hat{r}} \hat{r}.
\]

(13)

Table 4. \( \frac{\partial \hat{\alpha}}{\partial \hat{r}} \) used in Gradient Computing.

<table>
<thead>
<tr>
<th>Perspective</th>
<th>Para</th>
<th>Hyper</th>
<th>Equiangular</th>
<th>Stereographic</th>
<th>Orthogonal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Catadioptric</td>
<td>Catadioptric</td>
<td>Fisheye</td>
<td>Fisheye</td>
<td>Fisheye</td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>( \frac{1}{\sqrt{f^2 + f^2}} )</td>
<td>( \frac{\cot \chi}{\sqrt{f^2 + f^2}} )</td>
<td>( \frac{1}{2} \left( 1 - \frac{f}{\hat{r}} \cot \frac{\hat{r}}{2} + \cot^2 \frac{\hat{r}}{2} \right) )</td>
<td>( \frac{1}{\sqrt{f^2 - f^2}} )</td>
<td></td>
</tr>
</tbody>
</table>

The gradient of the line-image is used for several purposes. One of them is to define a more accurate threshold for the algebraic distance criterion. Having the two defining points of the curve and the gradient in each point we compute the coordinates of a point located to a metric distance from the curve. Then we compute the algebraic distance \( d \) of this point using the expression (10). As \( d \) is monotone the obtained distance could be used as threshold for the algebraic distance. The analytical gradient of the line-image is also used as additional criterion in the voting process. Having the orientation of the gradient in each point of the curve we can compute the angular distance between this and the gradient obtained in the Canny edge detection. Finally, the gradient is used to extract adjacent regions around the fitted segment of the curve.

![Fig. 6. Line-image extraction example on simulated images: (a) Hypercatadioptric System. (b) Fisheye Equiangular. (c) Fisheye Orthogonal.](image-url)
4.3 Robust Extraction on the Image

Line-image extraction can be explained as follows. First we detect the edges using the Canny algorithm and stored them in connected components. From Canny algorithm we also obtain the gradient of each pixel of the image.

For each component we launch a RANSAC algorithm to robustly extract line-images. Two points of the connected component are selected randomly. With these two points a line-image is computed using the two-points line-image definition presented above (Section 4.1). Two distances to the curve are computed from the rest of points of the connected component. First distance is algebraic distance shown in section 3.2. The second distance is an angular distance between the gradient in each point computed from the line-image (Section 4.2) and the gradient computed by image processing in the Canny edge detection. Points with both distances smaller than a threshold vote for this line-image. The candidate which collects more votes is selected as the best fit. Notice that,
this proposal assumes that a component contains at least a line-image. When a component is the projection of another shape (e.g. a circle on a planar surface) the algorithm does not fit the whole boundary. Instead of that, the algorithm extracts the line-image which better fits the given component. Once the line is fitted we extract the adjacent region surrounding the curve. Given a region thickness, the analytical gradient of each point of the segment is used to obtain the coordinates of the region. This image regions can be used for computing local-descriptors in order to perform a line-matching approach.

5 Experiments

We have tested the line extraction method using synthetic and real images. The synthetic images have been generated via a Matlab simulation. From each normal vector \( \mathbf{n} \) we compute the points of the intersection between the plane \( \Pi \) and the sphere. Then points are projected using the corresponding projection model. We have generated images for hypercatadioptric, equiangular or orthographic fisheye systems with a resolution of 1024x768 pixels. In Fig. 6(a-c) we show the line-images extracted for three simulated images in hypercatadioptric, equiangular and orthogonal systems. We show how the shape of the extracted curves are quite different depending on the projection model. Two different omnidirectional systems have been used to acquire the real images. The real dioptric images have been taken with an iPhone 4S camera with a commercial equiangular fisheye \(^4\) with a resolution of 3264x2498 pixels. The real hypercatadioptric images have been acquired with a firewire camera with an hyperbolic mirror and a resolution of 1024x768 pixels. In Fig. 7(a) we show the behaviour of the line-image extractor with a real equiangular image. In Fig. 7(b) we show the same image in which segments and its adjacent regions have been extracted. This test has been repeated for a hypercatadioptric image in Fig. 7(c-d).

6 Conclusions

We have presented a framework to deal with line-projection in any radially symmetric central projection system. This framework allows to perform line algorithms valid for different class of omnidirectional systems. Working on the image allow us to extract the adjoining regions surrounding each line-image. This image-regions can be used to compute region-based local descriptors in omnidirectional images. Notice that the extracted regions conserve invariance to orientation. Experimental results have been showed for catadioptric, equiangular-fisheye and orthogonal-fisheye models and the framework can be easily extended to other projection systems if they can expressed in the form \( \phi = h^{-1}(\hat{r}) \).

Acknowledgement. This work was supported by the Spanish projects VISPA DPI2009-14664-C02-01, VINEA DPI2012-31781, DGA-FSE(T04) and FEDER funds. First author was supported by the FPU program AP2010-3849.

References