Line extraction in uncalibrated central images with revolution symmetry

Jesus Bermudez-Cameo  
bermudez@unizar.es  
Gonzalo Lopez-Nicolas  
gonlopez@unizar.es  
Jose J. Guerrero  
josechu.guerrero@unizar.es  

Instituto de Investigación en Ingeniería de Aragón (I3A)  
Departamento de Informática e Ingeniería de Sistemas  
Universidad de Zaragoza, Spain

Abstract

In omnidirectional cameras, straight lines in the scene are projected onto curves called line-images. The shape of these curves is strongly dependent of the particular camera configuration. The great diversity of omnidirectional camera systems makes harder the line-image extraction in a general way. Therefore, it is difficult to design uncalibrated general approaches, and existing methods to extract lines in omnidirectional images require the camera calibration. In this paper, we present a novel method to extract line-images in uncalibrated images which is valid for radially symmetric central systems. In our proposal, the distortion function is analytically solved for different types of camera systems, dioptric or catadioptric. We present the unified line-image constraints to extract the projection plane of each line and main calibration parameter of the camera from a single line-image. The use of gradient-based information allows computing both from a minimum of two image points. This scheme is used in a line-image extraction algorithm to obtain lines from uncalibrated omnidirectional images without any assumption about the scene. The algorithm is evaluated with synthetic and real images showing good performance.

1 Introduction

Lines and their projections have been widely used as feature for 3D reconstruction and vision recognition. The projection of a straight line on the image is in general a curve called line-image. In conventional perspective cameras 3D straight lines are projected onto 2D straight lines so each line on the image can be recovered directly from 2 points of the line. When the projection model is not linear line-images are curves whose curvature depends on the acquisition system. Therefore the calibration of the system is needed to define this curve from 2 points. However the constraint of lying on a line holds even if camera calibration is unknown [8] and the line-image can be extracted without calibration.

In most approaches to extract lines from central omnidirectional images [9, 10] camera calibration is used to back project the image points to normalized rays lying on a unitary sphere rounding the viewpoint. In this space the relationship of the line projection becomes linear and classical approaches can be used. In the particular case of catadioptric systems...
the curve defining a line-image is a conic. Other approaches extract these conics directly on the images. In \cite{Bermudez-Cameo} conics are extracted on the image using the calibration of the camera. In \cite{Bermudez-Cameo} conics are computed using classical conic fitting approaches and after that calibration of the camera is estimated.

In general, the curve for a line-image can be extracted adding more constraints because the distortion in the line-image is induced by the non linear projection \cite{Bermudez-Cameo, Bermudez-Cameo, Bermudez-Cameo}. The location of more edge points lying on the curve can be used as additional constraints. The gradient orientation of these edge points could also be used if appropriate formulation is developed. As calibration of the camera is involved in the shape of the curves, both calibration parameters and line-images can be extracted simultaneously. That means that when line-images are automatically extracted from uncalibrated omnidirectional images self-calibration is also estimated. In \cite{Bermudez-Cameo} a line based self-calibration approach using a radially symmetric distortion model is presented. They encode the distortion function using a polynomial and a discrete representation. However, in our proposal the distortion function is analytically solved for each projection model reducing the number of parameters to encode it. Besides that, instead of calibration, our work is focused in line and segments extraction and since we encode the calibration in a single parameter we reduce considerably the iterations needed in a robust extraction.

In this paper we present a method for line extraction in uncalibrated images which is valid for central dioptric and catadioptric cameras with symmetry of revolution. This extraction allows us to segment the collection of edges corresponding to line-images automatically. For this, we exploit two classes of line-image constraints: point’s location and point’s gradient constraint. From the extracted line-image we also obtain the corresponding projecting planes defining the line-constraint on the 3D space, since main calibration information is implicitly computed. In our method we do not assume any restriction in the orientation of 3D lines. The input is a single image in which the existence of line-images is assumed and the result of our proposal is a set of points groups lying on different line-images, the corresponding projection planes and the main calibration parameter of the system.

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

In this section we introduce the line-image description for central systems with revolution symmetry presented by Bermudez-Cameo et al. in \cite{Bermudez-Cameo}. Assuming revolution symmetry exists, the spherical coordinates of a ray $\mathbf{x} = (\sin \phi \cos \varphi, \sin \phi \sin \varphi, \cos \phi)^T$ are related with the polar coordinates of the image points $\hat{\mathbf{x}} = (\hat{r}, \hat{\theta})$. Assuming that principal point is known, the image point coordinates $\hat{\mathbf{x}} = (\hat{x}, \hat{y})^T$ are referred to the principal point. The azimuth angle $\varphi$ is directly mapped to the coordinate $\hat{\theta} = \pm \varphi$ (the sign depends on the device: ’-‘ for catadioptric and ’+‘ for dioptric) and the relation between the radius of the image $\hat{r}$ and angle $\phi$ depends on the projection model. In Table 1 2\textsuperscript{nd} column we show the expression of the radius $\hat{r}$ in the symmetric distortion model based on the sphere model for catadioptric systems \cite{Bermudez-Cameo, Bermudez-Cameo} and particular projection models for dioptric systems \cite{Bermudez-Cameo, Bermudez-Cameo, Bermudez-Cameo} where $f$, $p$, and $\chi$ are calibration parameters in each corresponding projection model.

Let $\Pi = (n_x, n_y, n_z, 0)^T$ be a plane defined by a 3D line and the viewpoint of the system $O$. The projected line associated to the 3D line can be represented by $\mathbf{n} = (n_x, n_y, n_z)^T$. Points $\mathbf{X}$ lying in the 3D line are projected on points $\mathbf{x}$ which represent the direction of the projecting
The straight-line constraint on the image space

As introduced in section 2, in central systems the projection of a 3D line is the plane \( \Pi \) which is described on the image by constraint (1). This constraint is a curve which is in general non-linear. However the projecting plane is completely defined by two points and the viewpoint. That implies that the calibration of the system is embedded in the geometry of this curve. Therefore if the calibration of the system is known, the projecting plane normal \( \mathbf{n} \)

\[
d (\hat{\mathbf{x}}) = n_x \hat{x} + n_y \hat{y} + n_z \hat{\alpha} (\hat{r}) = 0 ,
\]

where \( \hat{\alpha} \) is a different expression for each camera system depending on the radius and the model parameters (see Table 1 3\(^\text{rd}\) column). Notice that \( d (\hat{\mathbf{x}}) \) is in general an algebraic distance which is zero when a point lies on the line. The normal direction to the line-image is described by the gradient of the homogeneous expression (1) which is expressed by

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

The term \( \frac{\partial \hat{\alpha}}{\partial \hat{r}} \) depends also on the camera model (see Table 1 4\(^\text{th}\) column).

Finally, notice that in central systems with symmetry of revolution any 3D line lying on the horizontal plane passing through the viewpoint \( \mathbf{n} = (0, 0, 1)^T \) is projected on the vanishing line. This line is always mapped in a circle centred on the principal point and with radius \( \hat{r}_{vl} \) (see Table 1 5\(^\text{th}\) column). This radius depends only on the system geometry and can be used to calibrate the camera system. In this case, equation (1) simplifies to \( \hat{\alpha} (\hat{r}) = 0 \).
can be recovered from two image points. However, when constraint (1) is satisfied the line-
image can be recovered even if camera calibration is unknown. In this section we present a
method for computing the projecting plane normal \( \mathbf{n} \) and the main calibration parameter from
a minimum of 3 image points lying on this line-image. Our proposal relies on \( \hat{r}_v \) as main
calibration parameter which is common to any central system with revolution symmetry and
depends only on the system geometry.

Given at least two points lying on a line-image described by (1) we can obtain the normal
\( \mathbf{n} \) by solving the homogeneous linear system

\[
\begin{pmatrix}
\hat{x}_i & \pm \hat{y}_i & -\hat{\alpha}_i
\end{pmatrix}
\mathbf{n} = 0 \quad \text{for} \quad i = 1, \ldots, n \quad \text{with} \quad n \geq 2. \tag{3}
\]

Notice that two points of the image back-projected to the sphere and the viewpoint are
enough to define the plane \( \Pi \). When the number of equations defining (3) is greater than two
we obtain a redundant system. Therefore when having three points lying on a line-image the
rank of the homogeneous matrix must be two. Imposing this condition we find a constraint
using three points of a line-image. The line-image constraint can be written as

\[
l_1 \hat{\alpha}_1 + l_2 \hat{\alpha}_2 + l_3 \hat{\alpha}_3 = 0, \tag{4}
\]

where \( l_1 = \hat{x}_2 \hat{y}_3 - \hat{x}_3 \hat{y}_2, \ l_2 = \hat{x}_3 \hat{y}_1 - \hat{x}_1 \hat{y}_3 \) and \( l_3 = \hat{x}_1 \hat{y}_2 - \hat{x}_2 \hat{y}_1 \).

With three points lying on the line-image and equation (4) it is possible to obtain the
main calibration parameter. In Table 2 we show the different expressions depending on the
system. In perspective systems it is not possible to extract any additional information. For
paracatadioptric and stereographic-fisheye systems the radius \( \hat{r}_v \) is computed directly. For
other dioptric systems the line-image constraint is solved by minimization.

In hypercatadioptric systems the focal distance of the camera is needed to compute the
main calibration parameter \( r_{vl} \). This is because two calibration parameters are involved and
they cannot be coupled. However, in practice it is possible to complete the extraction of the
line-image with a rough approximation of \( f \).

<table>
<thead>
<tr>
<th>Perspective</th>
<th>( (l_1 + l_2 + l_3) = 0 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Paracatadioptric</td>
<td>( \hat{r}_v = \sqrt\frac{l_1^2 + l_2^2 + l_3^2}{l_1 + l_2 + l_3} )</td>
</tr>
<tr>
<td>Hypercatadioptric</td>
<td>( \hat{r}_v = \sqrt\left(l_1 \sqrt{\frac{r_1^2 + r_2^2 + r_3^2}{l_1 + l_2 + l_3}} + r_2 \right)^2 - f^2 )</td>
</tr>
<tr>
<td>Equiangular-Fisheye</td>
<td>( l_1 \hat{r}_1 + l_2 \hat{r}_2 + l_3 \hat{r}_3 = 0 )</td>
</tr>
<tr>
<td>Stereographic-Fisheye</td>
<td>( \hat{r}_v = \sqrt\frac{l_1^2 + l_2^2 + l_3^2}{l_1 + l_2 + l_3} )</td>
</tr>
<tr>
<td>Orthogonal-Fisheye</td>
<td>( l_1 \left( \frac{\pi}{2} \hat{r}_1 \right) + l_2 \left( \frac{\pi}{2} \hat{r}_2 \right) + l_3 \hat{r}_3 = 0 )</td>
</tr>
<tr>
<td>Equisolid-Fisheye</td>
<td>( l_1 \left( \frac{\pi}{2} \hat{r}_1 \right) + l_2 \left( \frac{\pi}{2} \hat{r}_2 \right) + l_3 \hat{r}_3 = 0 )</td>
</tr>
</tbody>
</table>

Table 2: Three points line-image constraint for different central projection systems with
revolution symmetry.

\( f \), the radius of the sphere. When constraint (1) is not satisfied the line-

\( \mathbf{dr} \) is a dyadic.
4 Gradient-based constraints

In this section we present the line-image gradients for central systems with revolution symmetry, which gives additional constraints for line-image fitting. They allow to reduce the minimum number of defining points from three to two. Given an intensity image $I(x,y)$, its gradient $\nabla I(x,y) = (\nabla I_x, \nabla I_y)$ of each edge lying on a line-image is aligned with the gradient of the line-image (2). The dot product between the analytical gradient and the vector $\nabla I^T = (-\nabla I_y, \nabla I_x)^T$ describes the alignment between both gradients. This expression can be written as

$$-\nabla I_x n_x \pm \nabla I_y n_y + n_z \frac{\partial \hat{\alpha}}{\partial \hat{r}} \frac{1}{\hat{r}} (\hat{x} \nabla I_y - \hat{y} \nabla I_x) = 0,$$

and gives an additional constraint when solving a line-image. In particular, if knowing the system calibration and given the image gradient and a point lying on a line-image the normal $n$ is solved by the linear system

$$\begin{pmatrix} \hat{x} & \pm \hat{y} & -\hat{\alpha} \\ -\nabla I_y & \pm \nabla I_x & \frac{\partial \hat{\alpha}}{\partial \hat{r}} \frac{1}{\hat{r}} (\hat{x} \nabla I_y - \hat{y} \nabla I_x) \end{pmatrix} n = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

whose solution is

$$n = \begin{pmatrix} \nabla I_x \hat{\alpha} + \frac{\partial \hat{\alpha}}{\partial \hat{r}} \frac{\hat{y}}{\hat{r}} (\hat{x} \nabla I_y - \hat{y} \nabla I_x), & \nabla I_y \hat{\alpha} - \frac{\partial \hat{\alpha}}{\partial \hat{r}} \frac{\hat{x}}{\hat{r}} (\hat{x} \nabla I_y - \hat{y} \nabla I_x), & \hat{y} \nabla I_y + \hat{x} \nabla I_x \end{pmatrix}^T.$$

Then, using constraint (5) with a pair of points, the normal $n$ can be computed. With the proposed approach, when using (1) or (5) for fitting line-images we need to know the class of the central system to compute the corresponding value of $\hat{\alpha}$. An alternative is considering a linear approximation for $\hat{\alpha}(\hat{r})$ but generic for any system. Thus, we propose to approximate $\hat{\alpha}(\hat{r})$ by a linear function referenced to the radius of the vanishing line. Notice that in this case ($\hat{r} = \hat{r}_{vl}$) implies $\hat{\alpha} = 0$ (see Section 2), and simplifying the expression it becomes

$$\hat{\alpha}(\hat{r}) = \frac{\partial \hat{\alpha}(\hat{r}_{vl})}{\partial \hat{r}} (\hat{r} - \hat{r}_{vl}).$$

Using this approximation the system (6) can be expressed by

$$\begin{pmatrix} \hat{x} & \pm \hat{y} & -\hat{r} \\ -\nabla I_y & \pm \nabla I_x & \frac{\partial \hat{\alpha}(\hat{r}_{vl})}{\partial \hat{r}_{vl}} \hat{r}_{vl} \end{pmatrix} m = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

which can be solved from the coordinates and the gradient of two image points obtaining the vector $m = \lambda \begin{pmatrix} n_x, n_y, n_z \frac{\partial \hat{\alpha}(\hat{r}_{vl})}{\partial \hat{r}_{vl}} - n_z \frac{\partial \hat{\alpha}(\hat{r}_{vl})}{\partial \hat{r}_{vl}} \hat{r}_{vl} \end{pmatrix}^T$. Notice that in this vector $n_z$ and $\frac{\partial \hat{\alpha}(\hat{r}_{vl})}{\partial \hat{r}_{vl}}$ are coupled, however the main calibration parameter $\hat{r}_{vl}$ can be computed. With $\hat{r}_{vl}$, the value $\hat{\alpha}$ is defined (Table 1) and $n$ can be recovered from (3).

5 Uncalibrated Line-Image Extraction

In this section we propose the algorithm for line-image extraction from uncalibrated omnidirectional cameras. The algorithm relies on the constraints presented in Section 3 (based on a minimal set of 3 points). Alternatively, the algorithm is also developed for using gradient constraints, based on a minimal set of 2 points. The proposed algorithm is described next.
The input of the method are the edges and their gradients obtained from Canny detector. These edges are stored in connected components called boundaries and the gradient orientation is low-pass filtered to reduce noise. Then, a first splitting of the boundaries is done based on the orientation angle. The threshold has been chosen to minimize the number of splits and therefore a sub-boundary can contain more than one segment. A subset of these sub-boundaries is selected following a heuristic criterion to avoid line-images which do not contain self-calibration information (e.g. vertical lines). Sub-boundaries are sorted by the descriptor $\kappa = s_{\text{size}} \cdot \Delta \theta$ which is greater when the sub-boundary is best conditioned to extract calibration from it. $s_{\text{size}}$ is the size of the sub-boundary and $\Delta \theta$ is the angle covered by the sub-boundary rounding the principal point. Once the sub-boundaries are sorted a sub-set of them are selected. In particular, the sub-boundaries which cover the 50% of an accumulated histogram of $\kappa$ are chosen.

Once we have selected boundaries containing line-images well conditioned for self-calibration we use a RANSAC-based approach [9]. Points of a boundary are selected randomly to generate candidate line-images which are voted by the other points of the boundary. Two approaches are presented depending on the constraint considered:

**3-points $\hat{r}_{vl}$ estimation:** From the location coordinates of a minimum of three points of the boundary and the line-image constraints $\hat{r}_{vl}$ is estimated. Depending on the camera system the corresponding expression is used (Table 2).

**2-points $\hat{r}_{vl}$ estimation:** From the location coordinates of a minimum of two points and theirs gradients we can compute the vector $\mathbf{m}$ using (9). No assumption about the nature of the camera system is needed. Then, $\hat{r}_{vl}$ is computed from $\mathbf{m}$ by dividing its fourth component between the third.

From any of the both used approaches, the normal vector $\mathbf{n}$ is computed from the defining points and $\hat{r}_{vl}$ using the expression (3). The distance to determine if a point is lying on the line-image is the algebraic distance (1). The threshold for the distance $d(\hat{\mathbf{x}})$ is computed for each candidate. The gradient of the points used defines the normal direction to the curve. A fixed threshold in pixels is used to compute points located at certain metric distance to the curve. From these points a dynamic threshold is computed using the algebraic distance (1).
Each point of the boundary whose distance is minor than the dynamic threshold gives a vote for the candidate. The most voted candidate is selected as best fit.

From the previous procedure we obtain for each line-image, the points lying on it, the corresponding normal vector, and a estimation of the calibration via the $\hat{r}_{vl}$ value. Finally, main calibration parameter and normals are refined in a Levenberg-Marquardt optimization process for the whole image.

## 6 Experiments

We present different experiments to validate the proposal and to compare the two approaches presented in section 5. The accuracy in line extraction and calibration is measured using synthetic images with known ground-truth. Finally, to show the behaviour of the method with real images two experiments are also presented.
6.1 Experiments with Synthetic Data

We have generated images for hypercatadioptric, equiangular-fisheye, stereographic-fisheye, orthogonal-fisheye and equisolid-fisheye systems with a resolution of 1024 × 768 pixels. The images are from a synthetic scenario composed by segments in a corridor. A measure of the accuracy of the proposal is the error in the estimation of the radius of the vanishing line $\hat{r}_{vl}$. Notice that, as this radius depends only on the calibration and it is a magnitude common to any central system, it is useful for comparing calibration estimations between central systems with different characteristics. In Fig. 2 we show the deviation of the estimation of the radius of the vanishing line $\hat{r}_{vl}$, which has been normalized to make it independent from the image size. $\hat{r}_{vl} = 1$ corresponds to the radius of a corner of the image (e.g. with a resolution of 1024x768 values in pixels of $\hat{r}_{vl}$ are divided by 640). When $\hat{r}_{vl} < 1$ the FOV of the system is greater than 180 degrees. Otherwise the FOV is smaller than 180 degrees. Fig. 2 (a) depicts the results using the 2-points-$\hat{r}_{vl}$ estimation and Fig. 2 (b) the results using the 3-points-$\hat{r}_{vl}$ estimation. We also show the error in degrees between the ground-truth normals describing each line projection and the estimated ones (Fig. 3).

We observe that, in general, accuracy in $\hat{r}_{vl}$ estimation decreases when the radius is greater than the image size (in figure values greater than one). That means that the smaller the FOV the lesser the accuracy of the algorithm. The extreme case is the perspective camera where $\hat{r}_{vl} = \infty$ and curvature of lines is independent from the focal distance. We also note that the hypercatadioptric case is more difficult to solve than others as a result of the multiple parameters involved. The accuracy decreases considerably in systems with FOV less than 180 degrees. However, the accuracy in the normal’s orientation does not increase when $\hat{r}_{vl}$ increases and even decreases. This is an effect of the set-up of the experiment in which the scene is far away from the camera when $\hat{r}_{vl}$ increases to assure the scene is seen by the camera. In Fig. 1 we show some examples of simulated images for different catadioptric and dioptric devices and the corresponding extracted line-images.

6.2 Experiments with Real Images

In order to show how the method works with real images some tests have been performed with catadioptric and fisheye cameras. The principal point has been coarsely estimated from the intersection of vertical lines (which are radial lines lying on the principal point) by setting the axis of symmetry parallel to these vertical lines. In Fig. 4 (a-b) we show an hypercatadioptric image in which line-images have been extracted using our proposal. The image has been acquired with a hypercatadioptric camera mounted on a helmet with a resolution of 1024 × 768 pixels. In Fig. 4 (c-d) we apply the same algorithm but using a fisheye-camera image. The image has been taken using an iPhone 4S camera (3264 × 2498 pixels) with a equiangular fisheye. Notice that line-images of parallel 3D lines intersect in the vanishing points of the image because line-images have been well extracted.

In the second experiment we have applied the algorithm to an image sequence taken with a camera in hand. The sequence has been acquired with a Nexus 4 camera using an equiangular fisheye with a resolution of 1920 × 1080 pixels. The video of the sequence with the extracted line-images has been attached as supplementary material. Line-images from video (see Fig. 5(a)) intersect in the vanishing points of the image (notice that we do not have imposed this constraint). Finally in Fig. 5(b) we show the estimation of the main calibration parameter $\hat{r}_{vl}$ along the sequence (notice that each frame is independently computed). The mean estimation for $\hat{r}_{vl}$ is 568.41 pixels and the standard deviation is 9 pixels.
Figure 4: Line extraction on uncalibrated real omni-images. (a-b) Hypercatadioptric, (c-d) Equiangular-Fisheye

Figure 5: Line extraction on uncalibrated sequence: (a) Frame 182, (b) $\hat{r}_{vl}$ estimation in blue, $\bar{r}_{vl}$ mean value in red.

7 Conclusion

We have presented a method to extract line-images from uncalibrated central images with revolution symmetry. We consider a general framework which encodes projection models for dioptric and catadioptric systems using a common main calibration parameter (the radius of the vanishing line $\hat{r}_{vl}$). Due to the relation between projection models and curvature of line-images an estimation of the main calibration parameter of the system is obtained simultaneously. We observed that line-images are correctly classified and the obtained normals are
accurate enough to be used in 3D computations. Nevertheless, in hypercatadioptric systems it is difficult to decouple calibration parameters. In this case the accuracy is acceptable when the FOV of the system is greater than 180 degrees. Accuracy differences between the two proposed approaches for computing $\hat{r}_{vl}$ are not significant in spite of the 2-points approach being computationally lighter.

### Acknowledgement

This work was supported by the Spanish project VINEA DPI2012-31781, DGA-FSE(T04) and FEDER funds. First author was supported by the FPU program AP2010-3849.

### References


