Computer Vision on Rolling Shutter Cameras

PART II: Rolling Shutter Geometry

Per-Erik Forssén, Erik Ringaby, Johan Hedborg

Computer Vision Laboratory
Dept. of Electrical Engineering
Linköping University
# Tutorial overview

<table>
<thead>
<tr>
<th>Time</th>
<th>Topic</th>
<th>Presenter</th>
</tr>
</thead>
<tbody>
<tr>
<td>1:30–2:00pm</td>
<td>Introduction</td>
<td>Per-Erik</td>
</tr>
<tr>
<td>2:00–2:15pm</td>
<td><strong>Rolling Shutter Geometry</strong></td>
<td>Per-Erik</td>
</tr>
<tr>
<td>2:15–3:00pm</td>
<td>Rectification and Stabilisation</td>
<td>Erik</td>
</tr>
<tr>
<td>3:00–3:30pm</td>
<td>Break</td>
<td></td>
</tr>
<tr>
<td>3:30–3:45pm</td>
<td>Rolling Shutter and the Kinect</td>
<td>Erik</td>
</tr>
<tr>
<td>3:45–4:30pm</td>
<td>Structure from Motion</td>
<td>Johan</td>
</tr>
</tbody>
</table>
Projective Geometry

Textbook material

- Hartley & Zisserman
  Multiple View Geometry
  2nd ed 2004

- Faugeras
  Three-Dimensional Computer Vision
  1993
The pin-hole camera

Pin-hole camera model
The pin-hole camera

Pin-hole camera model

A brightly illuminated scene will be projected onto a wall opposite of the pin-hole.

The image is rotated 180°.
The pin-hole camera

From similar triangles we get:

\[ x = f \frac{X}{Z} \quad y = f \frac{Y}{Z} \]

\[ \gamma \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \]
The pin-hole camera

More generally we write:

\[ \gamma \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \]

- \( \gamma \) - projection of optical centre
- \( f \) - focal length, \( s \) - skew, \( a \) - aspect ratio,
- \( c = [c_x, c_y] \) - projection of optical centre
The pin-hole camera

\[ \gamma \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & s & c_x \\ 0 & a_f & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} \]

\[ x \sim K\tilde{X} \]

**Motivation:**

- f - focal length, s - skew, a - aspect ratio,
- \( c = [c_x \ c_y] \) - projection of optical centre

Image Plane

Image Grid

Optical Centre
The pin-hole camera

Projection of 3D points $X_c$ in the camera coordinate system:

$$x \sim K\tilde{X}_c$$

In order to relate several camera poses, we need to use a common world coordinate system (WCS):

$$\tilde{X}_c = R^T(\tilde{X}_w - d) \quad \Rightarrow \quad x \sim KR^T(\tilde{X}_w - d)$$

$d$ is a translation of the origin, and $R$ is a rotation
The pin-hole camera

We can simplify this to a single projection operation on the 3D points \( X_w \):

\[
x \sim K R \tilde{X}_w - d \quad \Rightarrow \quad x \sim P \tilde{X}_w
\]

where \( P \) is a 3x4 matrix, and

\[
X_w = \begin{bmatrix} \tilde{X}_w^T & 1 \end{bmatrix}^T = [X \ Y \ Z \ 1]^T
\]
The pin-hole camera

We can simplify this to a single projection operation on the 3D points \( \mathbf{X}_w \):

\[
x \sim \mathbf{K} \mathbf{R}^T (\mathbf{\bar{X}}_w - \mathbf{d}) \quad \Rightarrow \quad x \sim \mathbf{P} \mathbf{X}_w
\]

where \( \mathbf{P} \) is a 3x4 matrix, and

\[
\mathbf{X}_w = \begin{bmatrix} \mathbf{\bar{X}}_w^T & 1 \end{bmatrix}^T = \begin{bmatrix} X & Y & Z & 1 \end{bmatrix}^T
\]

This matrix \( \mathbf{P} \) has the explicit form:

\[
\mathbf{P} = \mathbf{K} \begin{bmatrix} \mathbf{R}^T \end{bmatrix} - \mathbf{R}^T \mathbf{d}
\]
The pin-hole camera

- We can simplify this to a single projection operation on the 3D points $X_w$:

  $$x \sim KR^T(\tilde{X}_w - d) \Rightarrow x \sim PX_w$$

- where $P$ is a 3x4 matrix, and

  $$X_w = \begin{bmatrix} \tilde{X}_w^T & 1 \end{bmatrix}^T = [X \ Y \ Z \ 1]^T$$

- This matrix $P$ has the explicit form:

  $$P = K \begin{bmatrix} R^T \mid - R^T d \end{bmatrix} = KR^T \begin{bmatrix} I \mid - d \end{bmatrix}$$
Rolling shutter model

Now recall the rolling shutter readout:

**Mechanical global shutter**

<table>
<thead>
<tr>
<th>Frame 1 Integration</th>
<th>Frame 1 Readout</th>
<th>Frame 2 Integration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Line 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Line 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Line 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Line 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Line n</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Electronic rolling shutter**

<table>
<thead>
<tr>
<th>Frame 1 Readout</th>
<th>Frame 2 Readout</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

---

CVPR 2012
Providence, Rhode Island
June 16-21, 2012

Linköping University
INSTITUTE OF TECHNOLOGY

13
Rolling shutter model

For a moving camera, projection in frame $k$ becomes:

$$x_k \sim KR_k^T [I| - d_k] X$$

$$x_k \sim KR(x_2)^T [I| - d(x_2)] X$$

**Mechanical global shutter**

<table>
<thead>
<tr>
<th>Frame 1</th>
<th>Frame 1</th>
<th>Frame 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Integration</td>
<td>Readout</td>
<td>Integration</td>
</tr>
<tr>
<td>Line 1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Line 2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Line 3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Line 4</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Line n</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Electronic rolling shutter**

<table>
<thead>
<tr>
<th>Frame 1</th>
<th>Frame 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Readout</td>
<td>Readout</td>
</tr>
<tr>
<td>Time</td>
<td>Time</td>
</tr>
</tbody>
</table>
Rolling shutter model

Overflow camera, projection in frame $k$ becomes:

$$x_k \sim KR_k^T [I - d_k] X$$

$$x_k \sim KR(x_2)^T [I - d(x_2)] X$$

Overflow the global shutter case, we have one pose $(R_k, d_k)$ per frame

Overflow the rolling shutter case, we instead get one pose $(R(x_2), d(x_2))$ per image row, $x_2$
Time coordinate

When interpolating the camera pose based on the image row, \( x_2 \) it is convenient to express time in number of rows, instead of seconds.

Recall that the frame period \( T \), is divided into readout time \( t_r \) and inter-frame delay \( t_d \).

\[
1/f = T = t_r + t_d
\]

\( t_r \) corresponds to number of image rows \( N_r \), and \( t_d \) corresponds to number of blank-rows \( N_b \).

\[
N_b = N_r t_d / t_r = N_r (t_r / f - 1)
\]
Triangulation is the process of estimating a 3D point $X$ from two projections $x_1$ and $x_2$. 

![Diagram of triangulation process](image)
Triangulation

For the two points, we have:

\[ x_1 \sim P_1 X \]
\[ x_2 \sim P_2 X \]
Triangulation

For the two points, we have:

\[ x_1 \sim P_1 X \]
\[ x_2 \sim P_2 X \]

Triangulation is typically solved by so called optimal triangulation [Hartley&Zisserman’04]

\[ X^* = \arg \min_X \left[ d^2(x_1, P_1 X) + d^2(x_2, P_2 X) \right] \]

The point \( X \) is sought, for which the squared re-projection error in both images is minimized.

There exists a closed form solution, that is found by solving a 3rd degree polynomial.
Rolling shutter triangulation

- If we generalize triangulation to a (moving) rolling shutter rig, we get:

\[ \mathbf{X}^* = \arg \min_{\mathbf{X}} [d^2(x_1, \mathbf{P}_1 \mathbf{X}(t_1)) + d^2(x_2, \mathbf{P}_2 \mathbf{X}(t_2))] \]

- This has an unique solution, if, and only if \( t_1 = t_2 \), which happens if the point is projected in both images at the same time.

- That is, when the two points have the same y-coordinate. (Very rare!)
Rolling shutter SfM

Suggestion from [Ait-Aider & Berry, ICCV’09]: Solve for triangulation of all points, and the object motion at the same time (structure-from-motion SfM).

The projection constraints for a correspondence now assumes the form:

\[ x_1 \sim K \begin{bmatrix} R(t_1) | d(t_1) \end{bmatrix} X \]
\[ x_2 \sim KR_2 \begin{bmatrix} R(t_2) | d_2 + d(t_2) \end{bmatrix} X \]

The set of all such triangulation constraints can uniquely define the solution, if we assume a parametric form for \( R(t) \) and \( d(t) \). The most simple one is a linear motion(6dof).
Degeneracy

The full optimization problem now looks like this:

\[
\{X^*_k\}^K_{k=1}, R, d = \arg\min_{\{X_k\}^K_{k=1}, R, d} \left[ \sum_{k=1}^{K} d^2(x_{1,k}, P_1 X_k) + d^2(x_{2,k}, P_2 X_k) \right]
\]

where

\[
P_1 = K \left[ R(t_1) | d(t_1) \right]
\]

\[
P_2 = K R_2 \left[ R(t_2) | d_2 + d(t_2) \right]
\]

This problem has a unique solution when the motion is different from a pure translation along the x-axis.

That is, when \( t_1 \neq t_2 \) for most points.

(Otherwise we never observe point motion.)
Degeneracy

For a translation parallel to the line between the optical centra, there is an equivalent stationary structure.

Illustration by Ait-Aider and Berry
Degeneracy

For a translation parallel to the line between the optical centra, there is an equivalent stationary structure

Illustration by Ait-Aider and Berry
Degeneracy

For a translation parallel to the line between the optical centra, there is an equivalent stationary structure.

Illustration by Ait-Aider and Berry
Degeneracy

- Structure and motion (SfM) from two frames is unstable for sideways motion, when both cameras have the same readout speed (or are the same).

- If one of the cameras has a global shutter, both structure and motion can be obtained [Ait-Aider&Berry ICCV’09]

- If multiple frames are used, rolling shutter structure from motion (SfM) becomes stable again [Hedborg et al. CVPR’12].
Related scanning cameras

Pushbroom [Gupta and Hartley PAMI’97]
Single line camera that is moving.
Related scanning cameras

- **Pushbroom [Gupta and Hartley PAMI’97]**
  Single line camera that is moving.

- **Work on scanning LIDARs, e.g. archeological reconstruction work by [Ikeuchi et al.]**

- **[Bosse ICRA’09]** e.g. rotating, and bouncing LIDARs.

Images from lab of Katsushi Ikeuchi
Related scanning cameras

Crossed slits [Zomet et al. PAMI’03]

Projection rays from 3D points to the image plane are defined as intersections of two "slit planes".
Related scanning cameras

- Crossed slits [Zomet et al. PAMI’03]

- Horizontal/vertical slits
- Pin-hole camera rendering
- Z rotation of vertical slit
- X rotation of vertical slit
Related scanning cameras

- Crossed slits [Zomet et al. PAMI’03]

sweeping push-broom mosaic  standard push-broom mosaic
Related scanning cameras

Crossed slits [Zomet et al. PAMI’03]

sweeping push-broom mosaic

standard push-broom mosaic

rendered image

pin-hole camera image
Related scanning cameras

- Crossed slits [Zomet et al. PAMI’03]

- [Geyer et al. OMNIVIS’05] demonstrate that a rolling shutter camera is equivalent to a crossed-slits camera for a pure translation parallel to the image plane. (but not in general)

- A crossed-slits camera can thus be seen as a special case of the rolling shutter camera.
Summary

- In rolling shutter geometry, the camera trajectory is best modelled as continuous.

- There is an ambiguity between structure and sideways motion for two-frame rolling-shutter geometry.

- Other types of scanning geometries (push-broom and moving LIDAR) do not have the temporal regularity of rolling shutter geometry.
References

- Hartley, Zisserman, "Multiple View Geometry for Computer Vision"
- Geyer, Meingast, Sastry, "Geometric Models of Rolling Shutter Cameras", OMIVIS’05
- Gupta, Hartley, "Linear Pushbroom Cameras", PAMI’97
- Ait-Aider, Berry, "Structure and Kinematics Triangulation from a Rolling Shutter Stereo Rig", ICCV’09
- Bosse, Zlot “Continuous 3D Scan-Matching with a Spinning 2D Laser”, ICRA’09
- Hedborg, Forssén, Felsberg, Ringaby, "Rolling Shutter Bundle Adjustment", CVPR’12