Research Log - Week 07

JeffGWood@mavs.uta.edu

July 28, 2016

June 26, 2016 Started reading Chapter 2 of [Hartley2004] [1] for information regarding *Homographices*.

Worked on graphics regarding $Epipolar\ constraint$ for inclusion in thesis document

June 27, 2016 Continued reading Chapter 2 of [Hartley2004] [1] containing information on Homographies for purpose(s) of deriving Fundamental matrix formula as well as understanding Horizontal rectification used for matching features along scanlines of images.

SUMMARY: Transformations of points in the image plane can be grouped into the following categories:

• Isometries (Denoted by \mathbf{H}_E): Transformations in \mathbb{P}_2 including translation and rotation (including composites of the two) that peserve Euclidean-distance. Transformations are of the form

$$\begin{bmatrix} \epsilon \cos(\theta) & -\sin(\theta) & t_x \\ \epsilon \sin(\theta) & \cos(\theta) & t_y \\ 0 & 0 & 1 \end{bmatrix}$$

where $\epsilon = \pm 1$. Angles are preserved if $\epsilon = 1$, else if $\epsilon = -1$ angles are reversed (reflection across an axis).

• Similarity (Denoted by \mathbf{H}_S): Transformations include translation, rotation, and scaling. Matrices are of the form

$$\begin{bmatrix} s\cos(\theta) & -s\sin(\theta) & t_x \\ s\sin(\theta) & s\cos(\theta) & t_y \\ 0 & 0 & 1 \end{bmatrix}$$

where s is the scaling factor. While *distances* are not preserved, the *ratio* of *distances* and *angles* are preserved.

• Affine (Denoted by \mathbf{H}_A): Transformations include all linear transformations of translation, rotation, scaling, and shearing. Matrices are of the form

$$\begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix}$$

• **Projective** (Denoted by \mathbf{H}_P): Transformations in \mathbb{P}_2 that are linear transformations in \mathbb{R}_3 . Matrices are of the form

$$\begin{bmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{bmatrix}$$

- Chapter 2: Projective Geometry:
 - Section 2.1: Planar Geometry:
 - Section 2.2: The 2D projective plane:

Lines in \mathbb{R}^2 are detailed by $\mathbf{l} = [a, b, c]^\intercal$ and points as $\mathbf{x} = [x, y, 1]^\intercal$ such that $\mathbf{l}^\intercal \cdot \mathbf{x} = a \cdot x + b \cdot y + 1 = 0$. Coordinates $\mathbf{x} = [x, y, 0]^\intercal$ with a 0 instead of 1 in the last place represent a *point at infinity* since they are the only points where $a \cdot x + b \cdot y + c \cdot 0 = a \cdot x + b \cdot y + c' \cdot 0$ for the two *parallel* lines of $\mathbf{l} = [a, b, c]^\intercal$ and $\mathbf{l}' = [a, b, c']^\intercal$

Cross product of points \mathbf{x} and \mathbf{x}' result in line l joining the two points (i.e. $\mathbf{x} \times \mathbf{x}' = l$). Cross product of lines l and l' result in point \mathbf{x} where intersection of two lines (i.e. $l \times l' = \mathbf{x}$).

Circles and ovals can be reprsented by a conic-matrix of the form

$$\begin{aligned} 0 &= \mathbf{x}^{\mathsf{T}} \cdot \mathbf{C} \cdot \mathbf{x} \\ &= \left[\begin{array}{ccc} x & y & 1 \end{array} \right] \cdot \left[\begin{array}{ccc} a & b/2 & d/2 \\ b/2 & c & e/2 \\ d/2 & e/2 & f \end{array} \right] \cdot \left[\begin{array}{c} x \\ y \\ 1 \end{array} \right] \\ &= a \cdot x^2 + b \cdot xy + c \cdot y^2 + d \cdot x + e \cdot y + f \cdot 1 \end{aligned}$$

- Section 2.3: Projective transformations:

Point \mathbf{x} on an image is mapped to point \mathbf{x}' via a homography \mathbf{H} , such that $\mathbf{x}' = \mathbf{H} \cdot \mathbf{x}$. Because a point \mathbf{x} lies on line \mathbf{l} if $\mathbf{l}^{\intercal} \cdot \mathbf{x} = 0$, then because

$$\begin{aligned} \mathbf{0} &= \mathbf{l}^{\mathsf{T}} \cdot \mathbf{x} \\ &= \mathbf{l}^{\mathsf{T}} \cdot \mathbf{H}^{-1} \cdot \mathbf{H} \cdot \mathbf{x} \\ &= \mathbf{l}^{\mathsf{T}} \cdot \mathbf{H}^{-1} \cdot \mathbf{x}' \end{aligned}$$

the point \mathbf{x}' lies on the line \mathbf{l}' defined by $\mathbf{l}'^{\mathsf{T}} = \mathbf{l}^{\mathsf{T}} \cdot \mathbf{H}^{-1}$, or $\mathbf{l}' = \mathbf{H}^{-\mathsf{T}} \cdot \mathbf{l}$. Therefore a homography that gives a *point-mapping* of $\mathbf{x}' = \mathbf{H} \cdot x$ has a corresponding *line-mapping* of $\mathbf{l}' = \mathbf{H}^{-\mathsf{T}} \cdot \mathbf{l}$.

Similarly, for a homography given by $\mathbf{x}' = \mathbf{H} \cdot \mathbf{x}$, the conic under the homography is given by

$$\begin{aligned} 0 &= \mathbf{x}^{\mathsf{T}} \cdot \mathbf{C} \cdot \mathbf{x} \\ &= (\mathbf{H}^{-1} \cdot \mathbf{x}')^{\mathsf{T}} \cdot \mathbf{C} \cdot (\mathbf{H}^{-1} \cdot \mathbf{x}') \\ &= \mathbf{x}'^{\mathsf{T}} \cdot \mathbf{H}^{-\mathsf{T}} \cdot \mathbf{C} \cdot \mathbf{H}^{-1} \cdot \mathbf{x}' \\ &= \mathbf{x}'^{\mathsf{T}} \cdot \mathbf{C}' \cdot \mathbf{x}' \end{aligned}$$

where $\mathbf{C}' = \mathbf{H}^{-\intercal} \cdot \mathbf{C} \cdot \mathbf{H}^{-1}$.

- Section 2.4: A hierarchy of transformations:

See entry from June 27, 2016.

- Chapter 6: Camera Models:
 - Section 6.1: Finite cameras:

Transformation from world-coordinate system \mathbf{x} to cameracoordinate system ${}^C\mathbf{x}$ is given by ${}^C\mathbf{x} = \mathbf{R} \cdot (\mathbf{x} - \mathbf{c})$. The Camera in world-space occurs at $\mathbf{x} = \mathbf{c}$. Camera-space has the camera located at ${}^C\mathbf{x} = 0$ and includes an image-plane at z = f. All rays intersect the image plane at z = f and converge on the origin ${}^C\mathbf{x} = 0$ which is known as the camera center. This results in points ${}^C\mathbf{x}$ in camera space being projected to points $\tilde{\mathbf{y}}$ in the image plane by means of the projection matrix \mathbf{P} such that

$$\mathbf{P} \cdot {}^{C}\tilde{\mathbf{x}} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} {}^{C}x_1 \\ {}^{C}x_2 \\ {}^{C}x_3 \\ 1 \end{bmatrix} = \begin{bmatrix} f \cdot {}^{C}x_1 \\ f \cdot {}^{C}x_2 \\ {}^{C}x_3 \end{bmatrix}$$
$$= {}^{C}x_3 \cdot \begin{bmatrix} f \cdot {}^{C}x_1/{}^{C}x_3 \\ f \cdot {}^{C}x_2/{}^{C}x_3 \end{bmatrix} = {}^{C}x_3 \cdot \tilde{\mathbf{y}}$$

This results in points containing infinitley large values of x_3 being mapped to the same principal point of $\mathbf{y}=0$ in the image plane. This assumes the principal point is always located in the image plane at $\mathbf{y}=0$. Projecting point $\tilde{\mathbf{x}}$ to the image plane with arbitrary principal point $\mathbf{p}=[p_x,p_y]$ requires modifying the projection matrix to include camera-specific parameters. The camera calibration matrix \mathbf{K} is given as

$$\mathbf{P} \cdot {}^{C}\tilde{\mathbf{x}} = \begin{bmatrix} f & 0 & p_{x} & 0 \\ 0 & f & p_{y} & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} {}^{C}x_{1} \\ {}^{C}x_{2} \\ {}^{C}x_{3} \\ 1 \end{bmatrix} = \begin{bmatrix} f \cdot {}^{C}x_{1} + p_{x} \cdot {}^{C}x_{3} \\ f \cdot {}^{C}x_{2} + p_{y} \cdot {}^{C}x_{3} \end{bmatrix}$$
$$= {}^{C}x_{3} \cdot \begin{bmatrix} f \cdot {}^{C}x_{1}/{}^{C}x_{3} + p_{x} \\ f \cdot {}^{C}x_{2}/{}^{C}x_{3} + p_{y} \end{bmatrix} = {}^{C}x_{3} \cdot \tilde{\mathbf{y}}$$

June 30, 2016 Question for Kamangar: On pages 162 and 244, how is the ray back-projected from \mathbf{x} by \mathbf{P} (where $\mathbf{x} = \mathbf{P}\mathbf{X}$ and $\mathbf{P} = \mathbf{K}[\mathbf{R}|\mathbf{t}]$) given by the formula $\mathbf{X}(\lambda) = \mathbf{P}^+\mathbf{x} + \lambda \mathbf{C}$? How is the formula derived?

July 1, 2016 Added section called **Points and Lines in the Image Plane** in the **Background** section.

References

[1] R. I. Hartley and A. Zisserman. *Multiple View Geometry in Computer Vision*. Cambridge University Press, ISBN: 0521540518, second edition, 2004.