

# Stereo Vision

Linda Tessens, IPI, TELIN

oktober 2005

## 1 Pinhole Cameramodel

We benaderen een reële camera door het pinhole model. Hierbij wordt een punt P in de ruimte geprojecteerd door de pinhole O op het beeldvlak  $\Pi$ . Uit Figuur 1 volgt:

$$\begin{cases} x' = f'x/z \\ y' = f'y/z \end{cases}$$

of in matrixnotatie:

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = 1/z \begin{bmatrix} f' & 0 & 0 \\ 0 & f' & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} \quad (1)$$

## 2 Homogene Coördinaten

Bij deze laatste vergelijking in matrixnotatie werd gebruik gemaakt van homogene coördinaten. Homogene coördinaten hebben als voordeel dat we er de bewerkingen die wij nodig hebben overzichtelijk en elegant mee kunnen noteren. Een punt in 2D

$$p = \begin{bmatrix} x \\ y \end{bmatrix}$$

wordt dan

$$p = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

en een punt in 3D analoog:

$$P = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

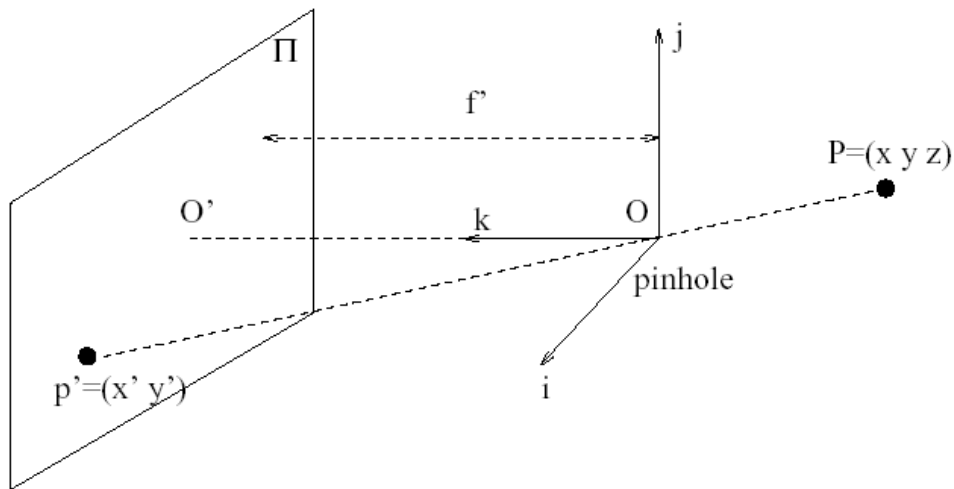


Figure 1: Perspectiefprojectie bij een pinhole camera [1].

wordt

$$P = \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

### 3 Geometrische Beeldvorming

#### 3.1 Interne cameraparameters

Het  $(x', y')$ -assenstelsel uit Figuur 1 heeft als centrum het snijpunt van de optische as van de camera met het beeldvlak. In een digitaal beeld wordt steeds gebruik gemaakt van een pixelcoördinaten-assenstelsel  $(u, v)$  met oorsprong in de linkerbovenhoek van het beeld en waarbij de horizontale  $u$ -as naar rechts wijst en de verticale  $v$ -as naar onderen. De  $(x', y')$ -coördinaten transformeren als volgt in de  $(u, v)$ -coördinaten

$$\begin{cases} u = k_x x' + x_0 \\ v = k_y y' + y_0 \end{cases}$$

waarbij  $k_x$  en  $k_y$  het aantal pixels per lengte-eenheid respectievelijk in de horizontale en verticale richting is.  $x_0$  en  $y_0$  zijn de coördinaten van het snijpunt van de optische as met het beeldvlak in het pixelcoördinaten-assenstelsel.

Samen met  $f$ , worden  $k_x$ ,  $k_y$ ,  $x_0$  en  $y_0$  de *interne calibratieparameters* van de camera genoemd.

### 3.2 Externe cameraparameters

Het punt P in Figuur 1 heeft de coördinaten  $(x, y, z)$  t.o.v. een assenstelsel verbonden met de camera (met als oorsprong de pinhole en met k-as loodrecht op het beeldvlak). De camera neemt een zekere positie in in de ruimte t.o.v. een wereldcoördinatenstelsel: ze is geroteerd en verschoven. Onderstel  $C(C_1, C_2, C_3)$  de coördinaten van de camera t.o.v. een wereldcoördinatenstelsel zijn, en

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

de rotatiematrix die de rotatie van het cameracoördinatenstelsel t.o.v. het wereldcoördinatenstelsel beschrijft. C en R worden de *externe calibratieparameters* van de camera genoemd.

### 3.3 Projectievergelijkingen

Een punt  $P_w$  met coördinaten  $(X_w, Y_w, Z_w)$  t.o.v. een wereldcoördinatenstelsel wordt afgebeeld op een punt met pixelcoördinaten  $(u, v)$ , en dit volgens volgende vergelijkingen:

$$\rho \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} k_x & 0 & x_0 \\ 0 & k_y & y_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} X_w - C_1 \\ Y_w - C_2 \\ Z_w - C_3 \end{bmatrix} \quad (2)$$

of korter

$$\rho p_{pix} = KR^t(P_w - C) \quad (3)$$

waarbij  $\rho$  een reële schaafactor verschillend van nul is. In homogene coördinaten kan deze vergelijking ook korter geschreven worden als

$$p_{pix} = \rho^{-1}K(R^t| - R^tC)P_w = TP_w. \quad (4)$$

## 4 3D Reconstructie [2]

We kunnen formule 3 ook omvormen tot

$$P_w = C + \rho RK^{-1}p_{pix} \quad (5)$$

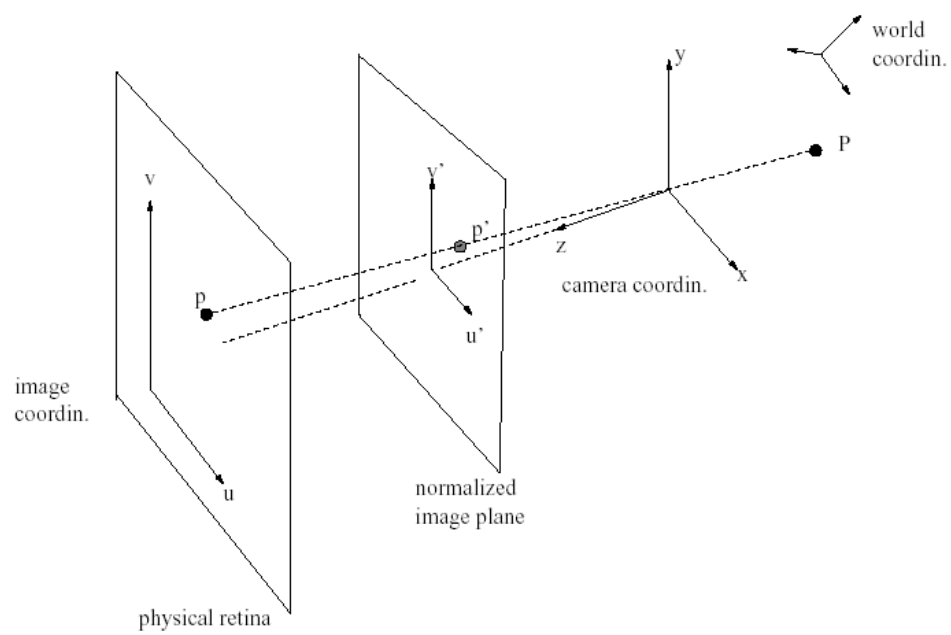


Figure 2: Geometrie van het beeldvormingsproces bij een perspectiefprojectie van een pinhole camera [1].

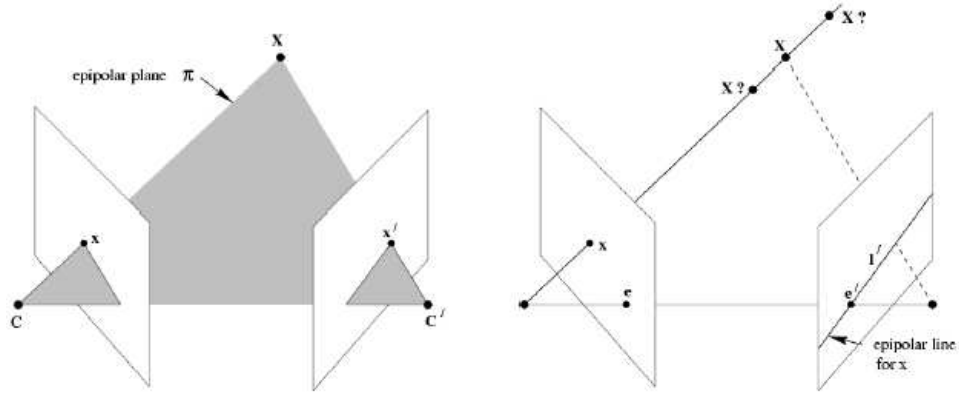


Figure 3: Epipolar line en epipole voor een stereo paar camera's [3].

of met andere woorden, het punt  $P_w$  ligt op de projectiestraal vanuit het cameracentrum  $C$  met richting  $RK^{-1}p_{pix}$ , waarbij de exacte positie op deze projectiestraal bepaald wordt door  $\rho$ .

Veronderstel nu dat een tweede camera het punt  $P_w$  eveneens waarneemt. Dan kan het punt  $P_w$  ook beschreven worden als

$$P_w = C' + \rho' R' K'^{-1} p'_{pix}. \quad (6)$$

5 en 6 leveren 6 vergelijkingen, voor 5 onbekenden ( $P_w$ ,  $\rho$  en  $\rho'$ ). Op deze manier kan men de ruimtelijke coördinaten van het punt  $P_w$  achterhalen. Door onnauwkeurigheden is het mogelijk dat dit overbepaald stelsel vergelijkingen geen oplossing heeft (dus dat de twee projectiestralen elkaar niet snijden). In dat geval kiest men voor  $P_w$  het midden van het kortste lijnstuk dat beide projectiestralen verbindt.

## 5 Epipolar Line

De projectiestraal die het punt  $P_w$  afbeeldt op  $p_{pix}$ , wordt in het beeldvlak van de tweede camera afgebeeld op een rechte, die de *epipolar line*  $l'$  van het punt  $p_{pix}$  genoemd wordt. Er kan aangetoond worden dat  $l'$  gegeven wordt door de vergelijking

$$l' = Fx$$

waarbij

$$F = [e']_x T' T^T (T T^T)^{-1}.$$

$$\begin{bmatrix} e_1 \\ e_2 \\ e_3 \end{bmatrix}_x = \begin{bmatrix} 0 & -e_3 & e_2 \\ e_3 & 0 & -e_1 \\ -e_2 & e_1 & 0 \end{bmatrix}$$
 is daarbij de matrixvoorstelling van het kruisproduct, en  $e'$  is de projectie van het cameracentrum  $C$  in het beeld van de tweede camera (de *epipole* van camera 1 in het beeld van camera 2;  $e' = T'C$ ).

## References

- [1] Ying Wu, “Camera model and image formation,” Tech. Rep., Electrical and Computer Engineering, Northwestern University, Evanston, IL 60208, 2005.
- [2] Seong-Wook Joo, “3d reconstruction - stereo,” Tech. Rep., Vision and AI Research Group, University of Maryland, 2004.
- [3] R. Hartley and A. Zisserman, *Multiple View Geometry in Computer Vision*, Cambridge University Press, 2000.