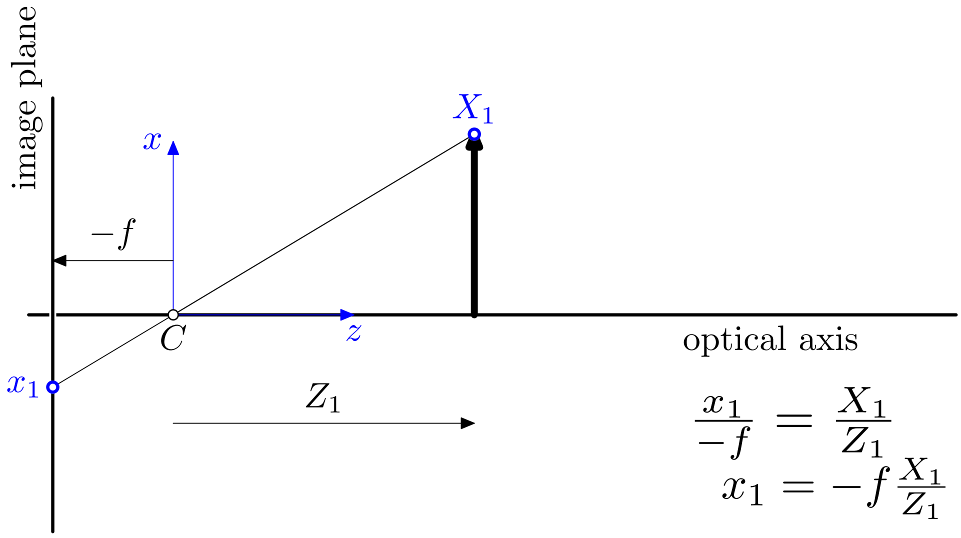
1D Pinhole camera projects 2D to 1D



$$\frac{x_1}{-f} = \frac{X_1}{Z_1}$$
$$x_1 = -f\,\frac{X_1}{Z_1}$$

- ideal lenses assumed
- coordinate system centered at C
- no $y - axis$
- similarity of projection triangles.

3D $\longrightarrow$ 2D Projection

We remember that: $\mathbf{x} = [\frac{fX}{Z}, \frac{fY}{Z}]^\top$

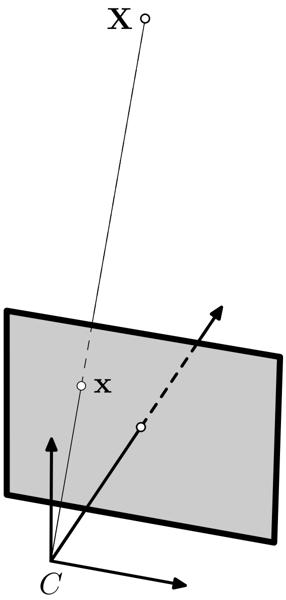
$$\begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix} \simeq \begin{bmatrix} fX \\ fY \\ Z \end{bmatrix}$$

$$\begin{bmatrix} \mathbf{x} \\ 1 \end{bmatrix} \simeq \begin{bmatrix} f & & & 0 \\ & f & & 0 \\ & & 1 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{X} \\ 1 \end{bmatrix}$$

Use the homegeneous coordinates[4]

$$\lambda_{[1\times1]}\mathbf{x}_{[3\times1]} = \mathtt{K}_{[3\times3]} \left[\ \mathtt{I}|\mathbf{0}\ \right] \mathbf{X}_{[4\times1]}$$

but . . .



[4] for the notation conventions, see the talk notes

Notation

scalars standard math italics, $f, \lambda \ldots$

vectors bold font $\mathbf{u}, \mathbf{x}, \mathbf{X}, \ldots$ are understood as *column* vectors. Row vectors are denoted by a symbol for transpose $\mathbf{x}^\top$.

matrices non-proportional font, $\mathtt{K}, \mathtt{R} \ldots$

$\simeq$ means equality upto a scale

Estimation of camera parameters—camera calibration

The goal: estimate the 3×4 camera projection matrix $\mathtt{P}$ and possibly the parameters of the non-linear distortion κ from images.

Assume a known projection $[u, v]^\top$ of a 3D point $\mathbf{X}$ with known coordinates

$$\begin{bmatrix} \lambda u \\ \lambda v \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{P}_1^\top \\ \mathbf{P}_2^\top \\ \mathbf{P}_3^\top \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

$$\frac{\lambda u}{\lambda} = \frac{\mathbf{P}_1^\top \mathbf{X}}{\mathbf{P}_3^\top \mathbf{X}} \quad \text{and} \quad \frac{\lambda v}{\lambda} = \frac{\mathbf{P}_2^\top \mathbf{X}}{\mathbf{P}_3^\top \mathbf{X}}$$

Re-arrange and assume[6] $\lambda \neq 0$ to get set of homegeneous equations

$$\begin{aligned} u\mathbf{X}^\top \mathbf{P}_3 - \mathbf{X}^\top \mathbf{P}_1 &= 0 \\ v\mathbf{X}^\top \mathbf{P}_3 - \mathbf{X}^\top \mathbf{P}_2 &= 0 \end{aligned}$$

[6] see some notes about $\lambda = 0$ in the talk notes

What does $\lambda = 0$ mean?

Point in the infinity hence, certainly out of the real image. $\lambda = 0$ makes pixel coordinates $u, v \to \infty$.

Decomposition of P into the calibration parameters

$$\mathtt{P} = \left[\ \mathtt{KR} \quad \mathtt{Kt}\ \right] \quad \text{and} \quad \mathbf{C} = -\mathtt{R}^{-1}\mathbf{t}$$

We know that R should be 3×3 orthonormal, and K upper triangular.

```
P = P./norm(P(3,1:3));

[K,R] = rq(P(:,1:3));
t = inv(K)*P(:,4);
C = -R'*t;
```

See the slide notes for more details.

```
P = P./norm(P(3,1:3));
```

The third row of the first 3×3 P submatrix should be actually equal to the third row of the rotation matrix R which is supposed to be *orthornomal*.

$$\mathtt{P}_{(:,1:3)} = \begin{bmatrix} \alpha_u & s & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \mathtt{R}$$

```
[K,R] = rq(P(:,1:3));
```

The RQ decomposition decomposes a matrix into a product of upper triangular and unitary matrix. It is less frequent than the more known QR decomposition. Implementatition of RQ decomposition may be found at numerous packages. Explanation may be read in [2, Appendix 4.1.1 (page 579)]

```
t = inv(K)*P(:,4);
C = -R'*t;
```

Recover the translation vector from the fourth column of P and compute the position of the camera center in the world coordinate system.

End

References

[1] Richard Hartley and Andrew Zisserman. *Multiple view geometry in computer vision.* Cambridge University, Cambridge, 2nd edition, 2003.