Computer Vision: Algorithms and Applications

Image Formation

Jing Luo | Megvii Tech Talk | Feb 2018

Reference: R. Szeliski. *Computer Vision: Algorithms and Applications*. 2010. 1.

1.
Geometric primitives and transformations

Geometric primitives

- 2D points

$$\boldsymbol{x} = \begin{bmatrix} x \\ y \end{bmatrix} \qquad \tilde{\boldsymbol{x}} = (\tilde{x}, \tilde{y}, \tilde{w}) = \tilde{w}(x, y, 1) = \tilde{w}\bar{\boldsymbol{x}}$$

- 2D lines

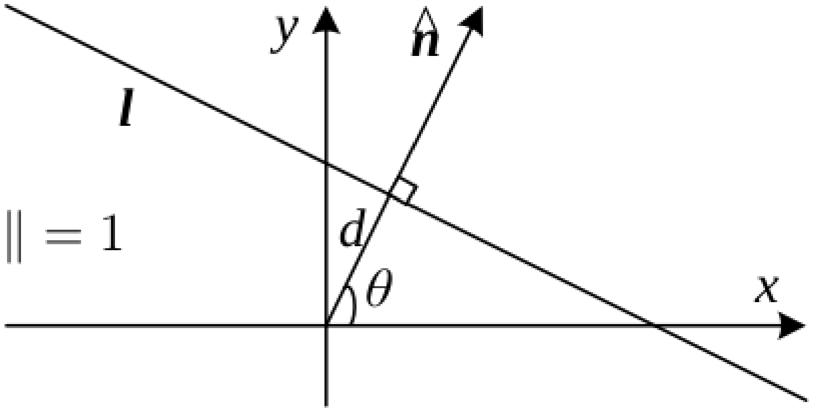
$$\tilde{\boldsymbol{l}} = (a, b, c)$$

$$\bar{\boldsymbol{x}} \cdot \tilde{\boldsymbol{l}} = ax + by + c = 0.$$

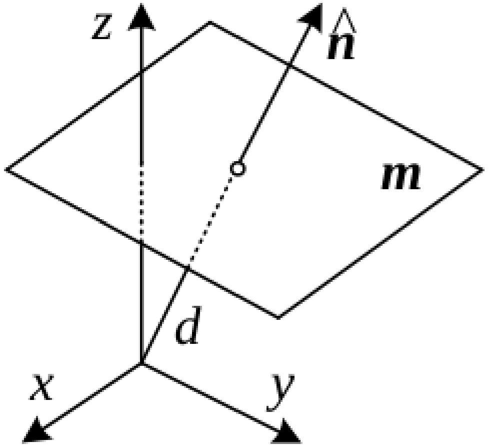
$$\boldsymbol{l} = (\hat{n}_x, \hat{n}_y, d) = (\hat{\boldsymbol{n}}, d) \text{ with } \|\hat{\boldsymbol{n}}\| = 1$$

$$\hat{\boldsymbol{n}} = (\hat{n}_x, \hat{n}_y) = (\cos\theta, \sin\theta)$$

polar coordinates (θ, d)



Geometric primitives



- Use homogeneous coordinates

  Intersection of two lines:

$$\tilde{\boldsymbol{x}} = \tilde{\boldsymbol{l}}_1 \times \tilde{\boldsymbol{l}}_2$$

  The line joining two points:

$$\tilde{\boldsymbol{l}} = \tilde{\boldsymbol{x}}_1 \times \tilde{\boldsymbol{x}}_2$$

- 3D points

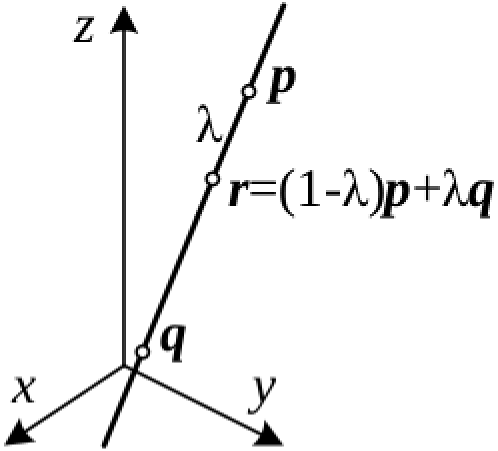
$$\boldsymbol{x} = (x, y, z) \in \mathcal{R}^3 \quad \tilde{\boldsymbol{x}} = (\tilde{x}, \tilde{y}, \tilde{z}, \tilde{w}) \in \mathcal{P}^3 \quad \bar{\boldsymbol{x}} = (x, y, z, 1)$$

- 3D planes

$$\bar{\boldsymbol{x}} \cdot \tilde{\boldsymbol{m}} = ax + by + cz + d = 0$$

Geometric primitives

- 3D lines



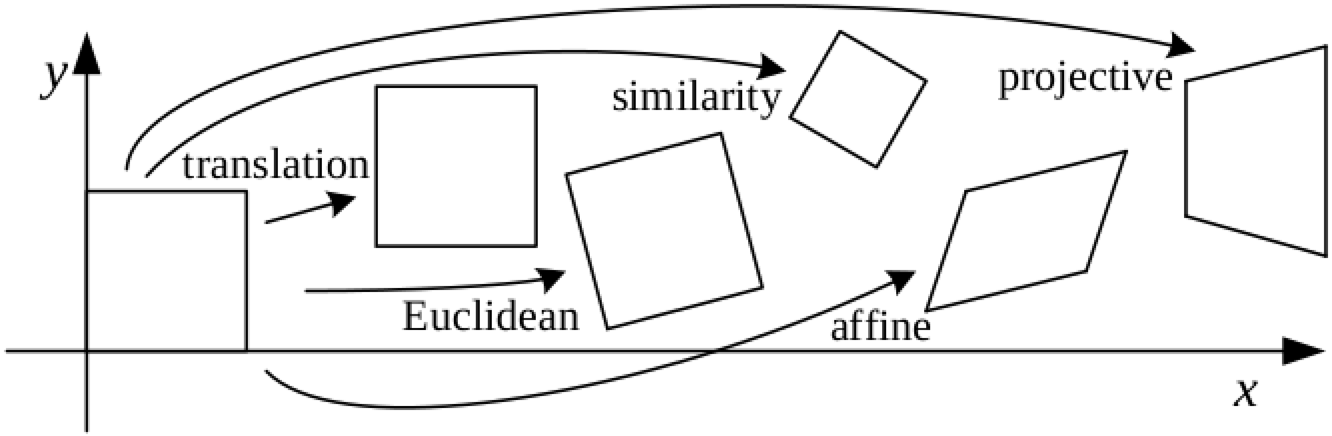
$$\boldsymbol{r} = (1-\lambda)\boldsymbol{p} + \lambda\boldsymbol{q}$$

$$\tilde{\boldsymbol{q}} = (\hat{d}_x, \hat{d}_y, \hat{d}_z, 0) = (\hat{\boldsymbol{d}}, 0)$$

$$\boldsymbol{r} = \boldsymbol{p} + \lambda\hat{\boldsymbol{d}}.$$

2D transformations

- Frequent 2D transformations



2D transformations

- Translation

$$\boldsymbol{x}' = \boldsymbol{x} + \boldsymbol{t}$$

$$\boldsymbol{x}' = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{t} \end{bmatrix} \bar{\boldsymbol{x}}$$

$$\bar{\boldsymbol{x}}' = \begin{bmatrix} \boldsymbol{I} & \boldsymbol{t} \\ \boldsymbol{0}^T & 1 \end{bmatrix} \bar{\boldsymbol{x}}$$

- Rotation + Translation

$$\boldsymbol{x}' = \boldsymbol{R}\boldsymbol{x} + \boldsymbol{t}$$

$$\boldsymbol{x}' = \begin{bmatrix} \boldsymbol{R} & \boldsymbol{t} \end{bmatrix} \bar{\boldsymbol{x}}$$

$$\boldsymbol{R} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$ 旋转 $\begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}$ 反射

2D transformations

- Scaled rotation

$$\boldsymbol{x}' = \begin{bmatrix} s\boldsymbol{R} & \boldsymbol{t} \end{bmatrix} \bar{\boldsymbol{x}} = \begin{bmatrix} a & -b & t_x \\ b & a & t_y \end{bmatrix} \bar{\boldsymbol{x}}$$

- Affine

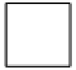
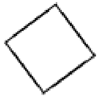
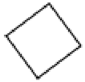
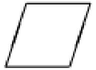
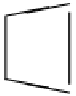
$$\boldsymbol{x}' = \begin{bmatrix} a_{00} & a_{01} & a_{02} \\ a_{10} & a_{11} & a_{12} \end{bmatrix} \bar{\boldsymbol{x}}$$

- Projective

$$\tilde{\boldsymbol{x}}' = \tilde{\boldsymbol{H}} \tilde{\boldsymbol{x}} \quad x' = \frac{h_{00}x + h_{01}y + h_{02}}{h_{20}x + h_{21}y + h_{22}} \text{ and } y' = \frac{h_{10}x + h_{11}y + h_{12}}{h_{20}x + h_{21}y + h_{22}}$$

2D transformations

- Hierarchy of 2D coordinate transformations

Transformation	Matrix	# DoF	Preserves	Icon
translation	$\begin{bmatrix} \boldsymbol{I} \mid \boldsymbol{t} \end{bmatrix}_{2\times 3}$	2	orientation	
rigid (Euclidean)	$\begin{bmatrix} \boldsymbol{R} \mid \boldsymbol{t} \end{bmatrix}_{2\times 3}$	3	lengths	
similarity	$\begin{bmatrix} s\boldsymbol{R} \mid \boldsymbol{t} \end{bmatrix}_{2\times 3}$	4	angles	
affine	$\begin{bmatrix} \boldsymbol{A} \end{bmatrix}_{2\times 3}$	6	parallelism	
projective	$\begin{bmatrix} \tilde{\boldsymbol{H}} \end{bmatrix}_{3\times 3}$	8	straight lines	

2D transformations

- Stretch/squash

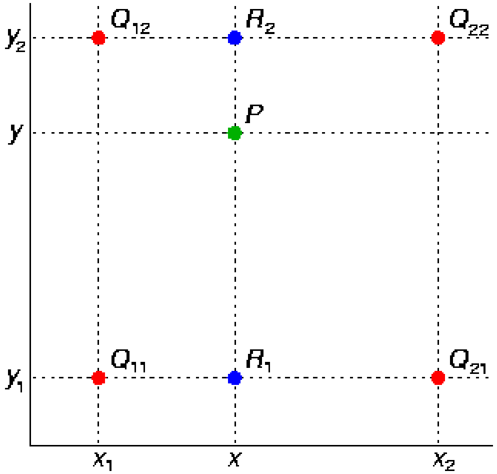
$$x' = s_x x + t_x$$
$$y' = s_y y + t_y$$

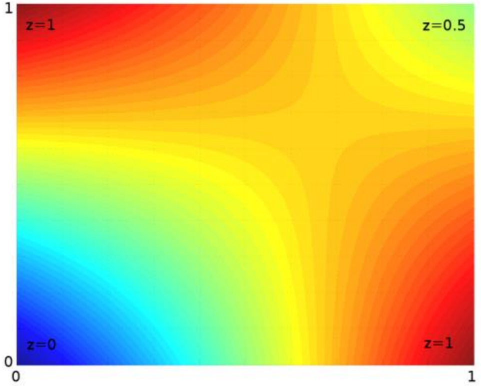
- Planar surface flow

$$x' = a_0 + a_1 x + a_2 y + a_6 x^2 + a_7 xy$$
$$y' = a_3 + a_4 x + a_5 y + a_7 x^2 + a_6 xy$$

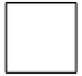
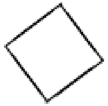
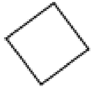
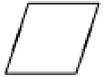
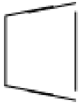
- Bilinear interpolation

$$x' = a_0 + a_1 x + a_2 y + a_6 xy$$
$$y' = a_3 + a_4 x + a_5 y + a_7 xy$$



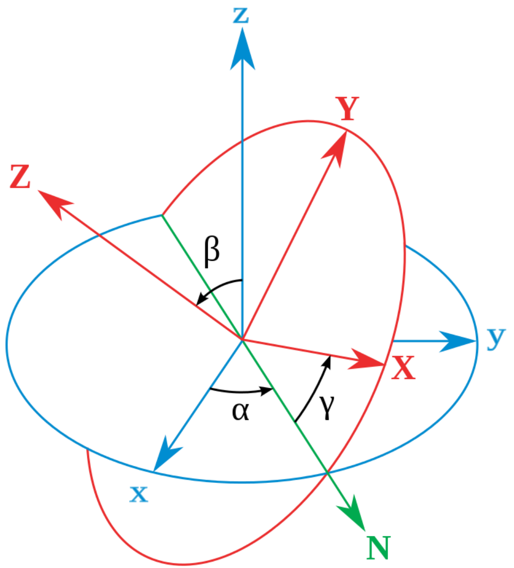


3D transformations

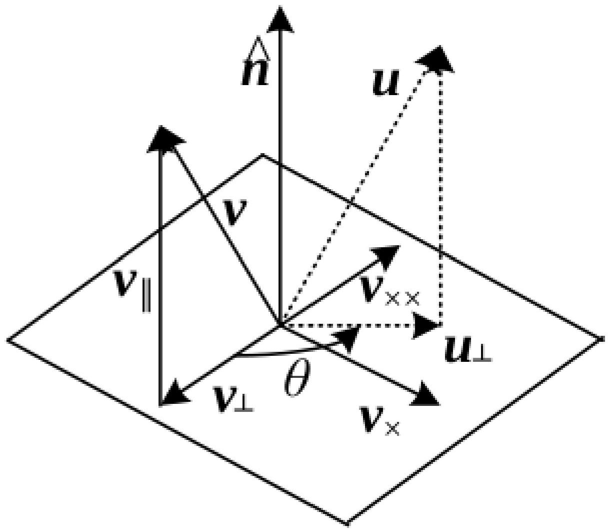
Transformation	Matrix	# DoF	Preserves	Icon
translation	$\begin{bmatrix} \boldsymbol{I} \mid \boldsymbol{t} \end{bmatrix}_{3\times 4}$	3	orientation	
rigid (Euclidean)	$\begin{bmatrix} \boldsymbol{R} \mid \boldsymbol{t} \end{bmatrix}_{3\times 4}$	6	lengths	
similarity	$\begin{bmatrix} s\boldsymbol{R} \mid \boldsymbol{t} \end{bmatrix}_{3\times 4}$	7	angles	
affine	$\begin{bmatrix} \boldsymbol{A} \end{bmatrix}_{3\times 4}$	12	parallelism	
projective	$\begin{bmatrix} \tilde{\boldsymbol{H}} \end{bmatrix}_{4\times 4}$	15	straight lines	

3D rotations

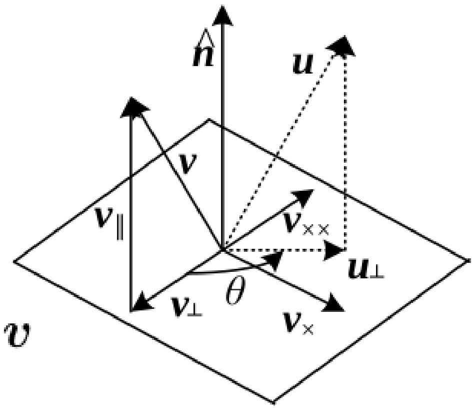
- Euler angles?



- Axis/angle



3D rotations



- Rodriguez's formula

$$\boldsymbol{u} = \boldsymbol{u}_\perp + \boldsymbol{v}_\parallel = (\boldsymbol{I} + \sin\theta[\hat{\boldsymbol{n}}]_\times + (1-\cos\theta)[\hat{\boldsymbol{n}}]_\times^2)\boldsymbol{v}$$

$$\boldsymbol{R}(\hat{\boldsymbol{n}}, \theta) = \boldsymbol{I} + \sin\theta[\hat{\boldsymbol{n}}]_\times + (1-\cos\theta)[\hat{\boldsymbol{n}}]_\times^2$$

$$\boldsymbol{R}(\boldsymbol{\omega}) \approx \boldsymbol{I} + \sin\theta[\hat{\boldsymbol{n}}]_\times \approx \boldsymbol{I} + [\theta\hat{\boldsymbol{n}}]_\times = \begin{bmatrix} 1 & -\omega_z & \omega_y \\ \omega_z & 1 & -\omega_x \\ -\omega_y & \omega_x & 1 \end{bmatrix}$$

3D rotations

- Unit quaternions

$$\begin{aligned}\boldsymbol{R}(\hat{\boldsymbol{n}},\theta) &= \boldsymbol{I}+\sin\theta[\hat{\boldsymbol{n}}]_\times+(1-\cos\theta)[\hat{\boldsymbol{n}}]_\times^2 \\ &= \boldsymbol{I}+2w[\boldsymbol{v}]_\times+2[\boldsymbol{v}]_\times^2.\end{aligned}$$

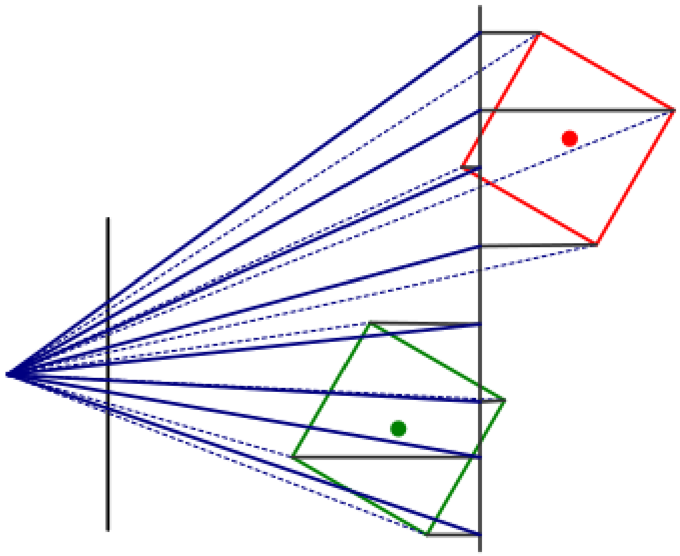
$$\boldsymbol{R}(\boldsymbol{q})=\begin{bmatrix}1-2(y^2+z^2) & 2(xy-zw) & 2(xz+yw) \\ 2(xy+zw) & 1-2(x^2+z^2) & 2(yz-xw) \\ 2(xz-yw) & 2(yz+xw) & 1-2(x^2+y^2)\end{bmatrix}$$

3D to 2D projections

- Orthography

$$\boldsymbol{x} = [\boldsymbol{I}_{2\times 2}|\boldsymbol{0}]\,\boldsymbol{p}$$

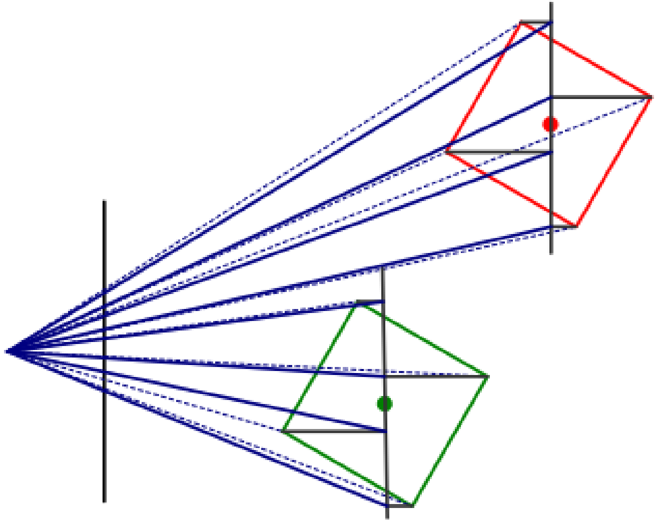
$$\tilde{\boldsymbol{x}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \tilde{\boldsymbol{p}}$$



3D to 2D projections

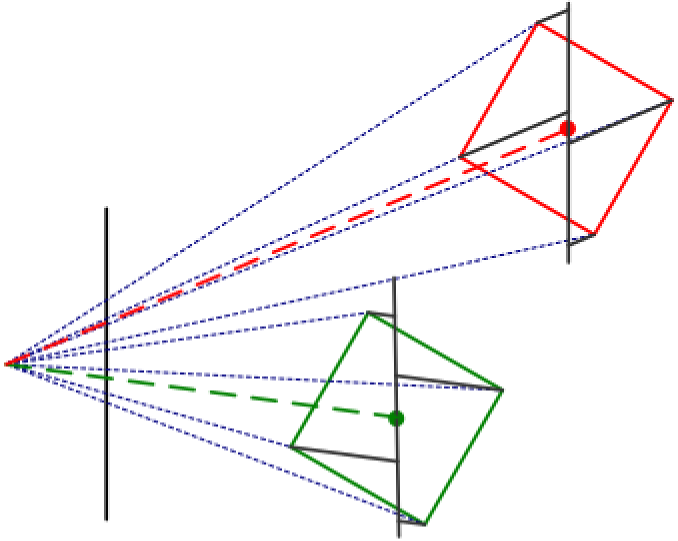
- Scaled orthography

$$x = [sI_{2\times 2}|0]\, p$$



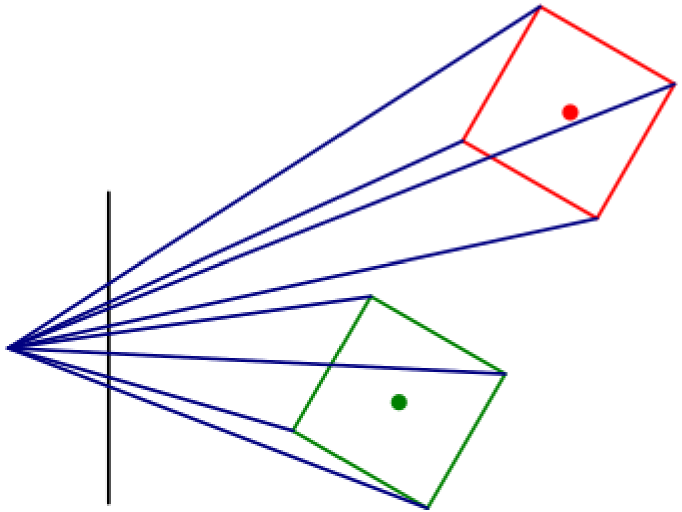
- Para-perspective

  Affine



3D to 2D projections

- 3D perspective



$$\bar{\boldsymbol{x}} = \mathcal{P}_z(\boldsymbol{p}) = \begin{bmatrix} x/z \\ y/z \\ 1 \end{bmatrix}$$

$$\tilde{\boldsymbol{x}} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \tilde{\boldsymbol{p}}$$

Lens distortions

- Straight lines in the world = straight lines in the image?
  Many wide-angle lenses have noticeable radial distortion

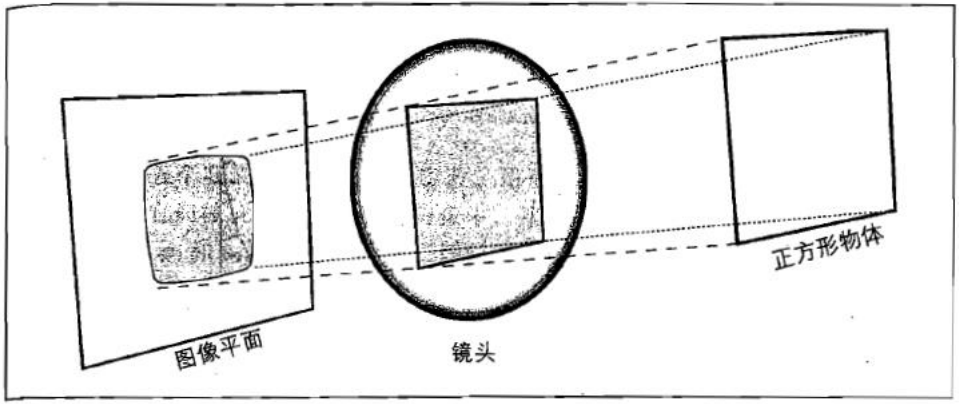


图 11-3：径向畸变。远离透镜中心的光线弯曲比靠近中心的严重。因此正方形的边在图像平面上为弯曲(即筒形畸变)

Lens distortions

(a) (b) (c)

Figure 2.13 Radial lens distortions: (a) barrel, (b) pincushion, and (c) fisheye. The fisheye image spans almost 180° from side-to-side.

Lens distortions

- Compensate for radio distortion

$$\begin{aligned} \hat{x}_c &= x_c(1 + \kappa_1 r_c^2 + \kappa_2 r_c^4) \\ \hat{y}_c &= y_c(1 + \kappa_1 r_c^2 + \kappa_2 r_c^4) \end{aligned}$$

2.
Photometric image formation

Lighting

- Point light

  Location, intensity, color spectrum $L(\lambda)$.

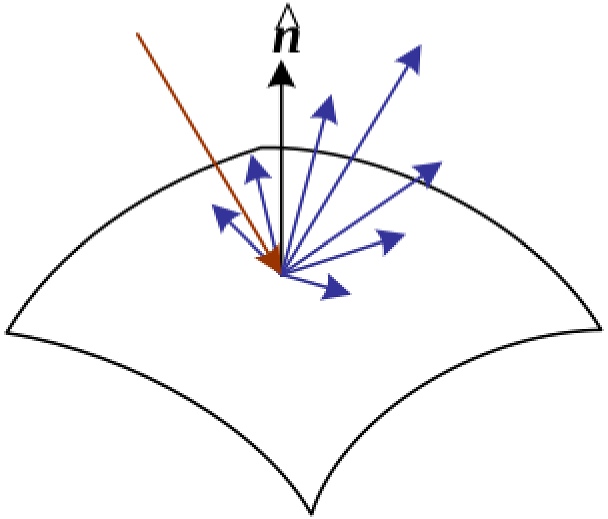
- Area light sources

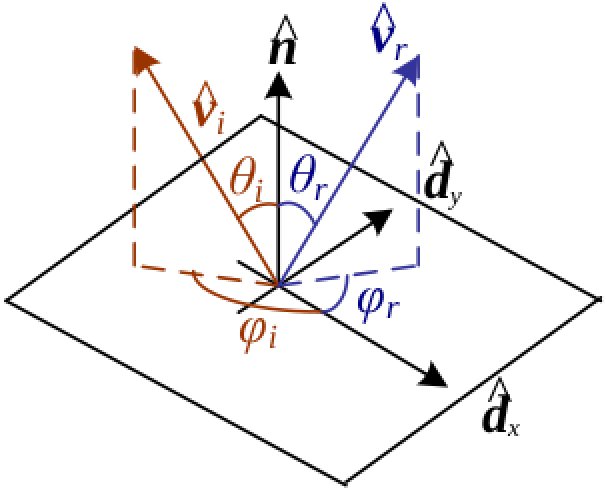
  A finite rectangular area emitting light equally in all directions.

  Environment map: maps incident light directions $\hat{v}$ to color values.

Reflectance and shading

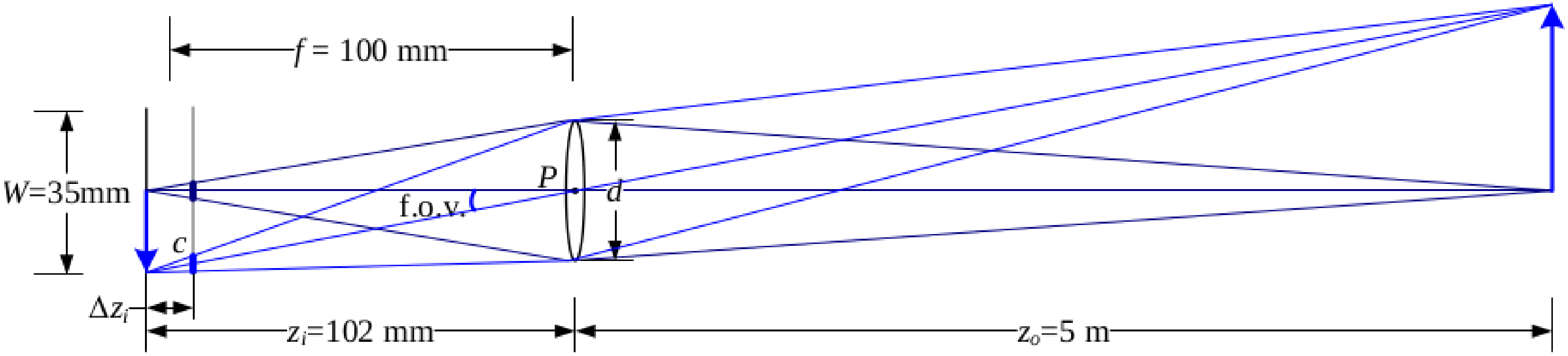
- Bidirectional Reflectance Distribution Function (BRDF)





Optics

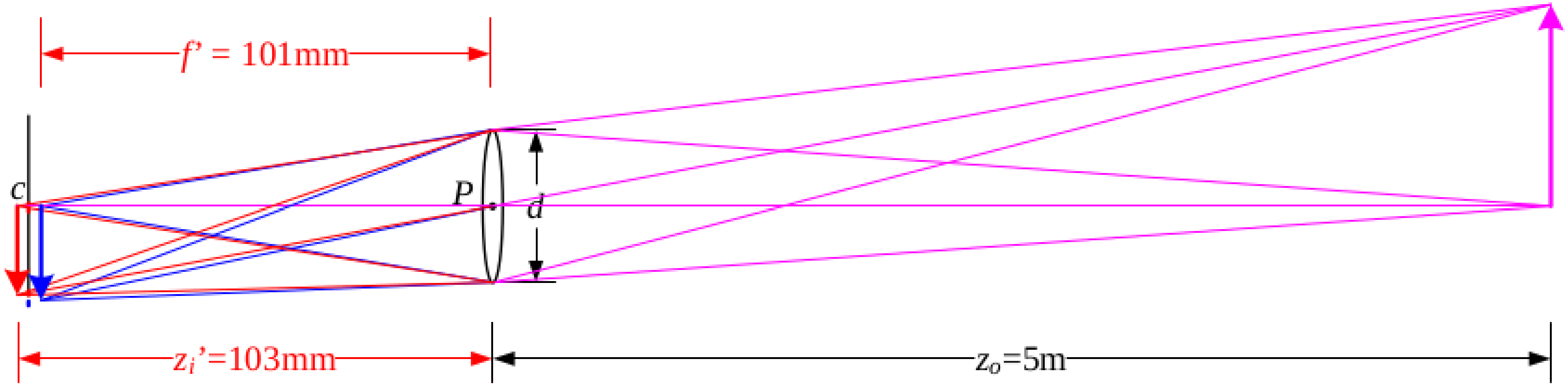
- Rule of optics



$$\frac{1}{z_o} + \frac{1}{z_i} = \frac{1}{f}$$

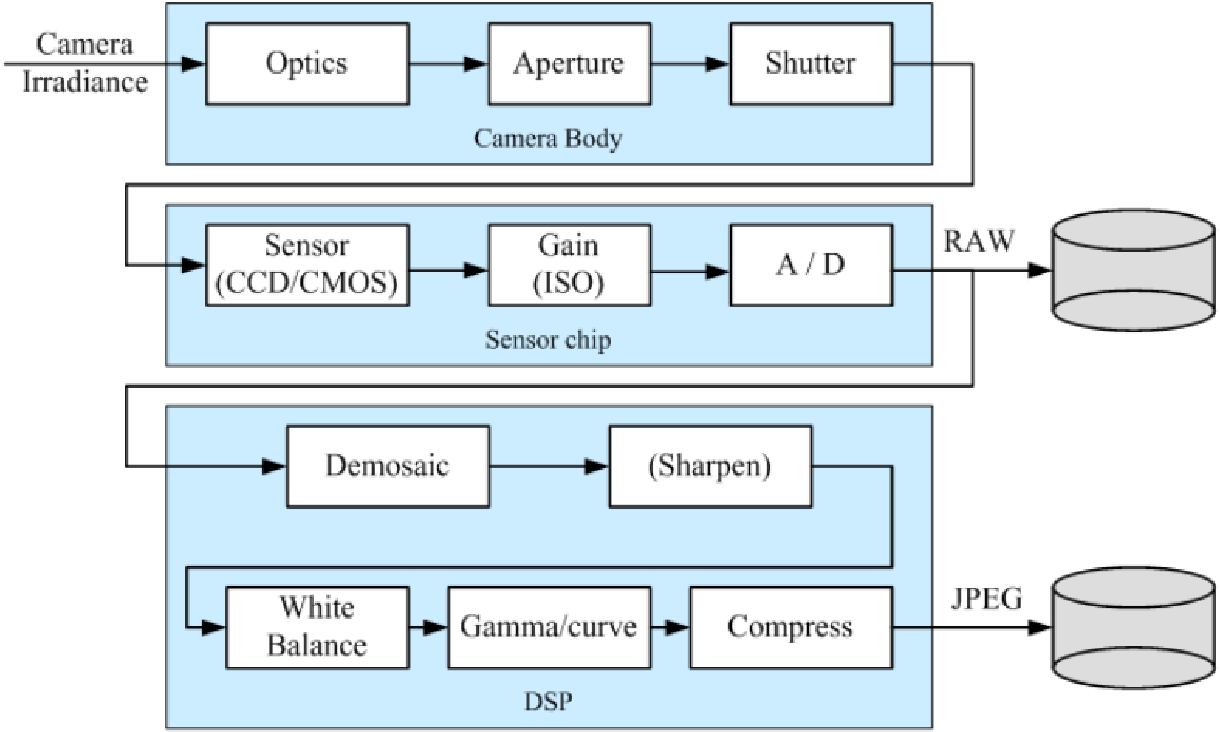
Optics

- Chromatic aberration



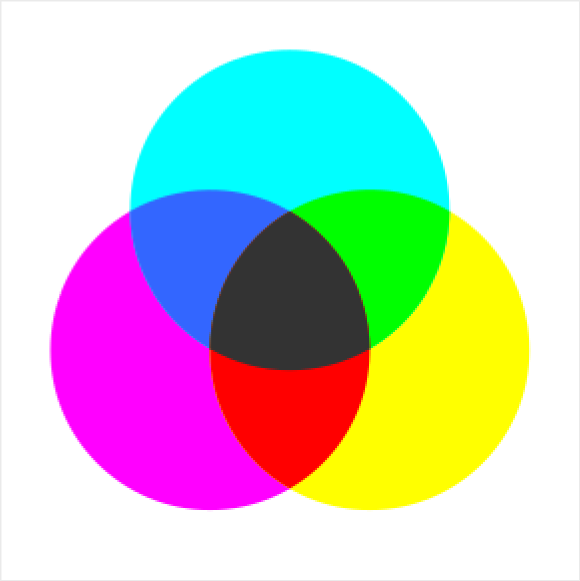
3.
The digital camera

Image sensing pipeline



Color

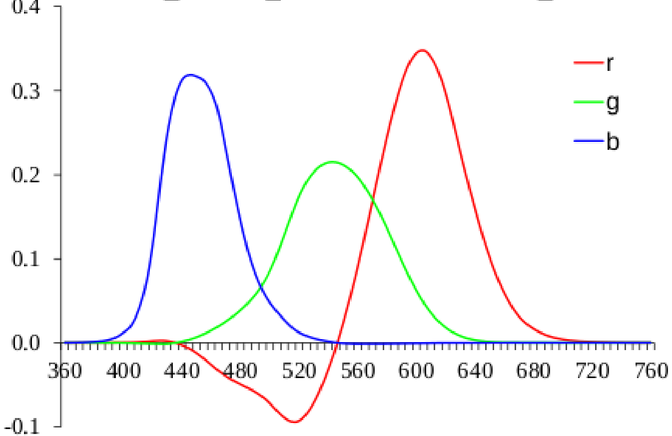
- Additive colors / subtractive colors

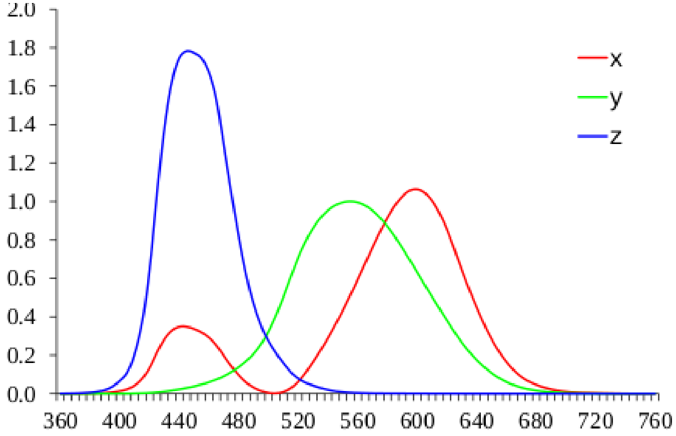


Color

- RGB / XYZ / LAB

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \frac{1}{0.17697} \begin{bmatrix} 0.49 & 0.31 & 0.20 \\ 0.17697 & 0.81240 & 0.01063 \\ 0.00 & 0.01 & 0.99 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$





Color filter arrays

- Bayer RGB pattern

G	R	G	R
B	G	B	G
G	R	G	R
B	G	B	G

rGb	Rgb	rGb	Rgb
rgB	rGb	rgB	rGb
rGb	Rgb	rGb	Rgb
rgB	rGb	rgB	rGb

Compression

- Discrete cosine transform (DCT)

  Both MPEG and JPEG use 8 × 8 DCT transforms.
- PSNR: quality of a compression algorithms

$$PSNR = 10 \log_{10} \frac{I_{\max}^2}{MSE} = 20 \log_{10} \frac{I_{\max}}{RMS}$$