

Corso di Visione Artificiale

The background features a large, faint, circular seal of Ca' Foscari University of Venice. The seal contains a central figure, likely a lion or a similar heraldic symbol, surrounded by Latin text. The top arc of the seal reads "CA' FOSCARI UNIVERSITAS" and the bottom arc reads "IN DOMO FOSCARI".

Camera

Samuel Rota Bulò

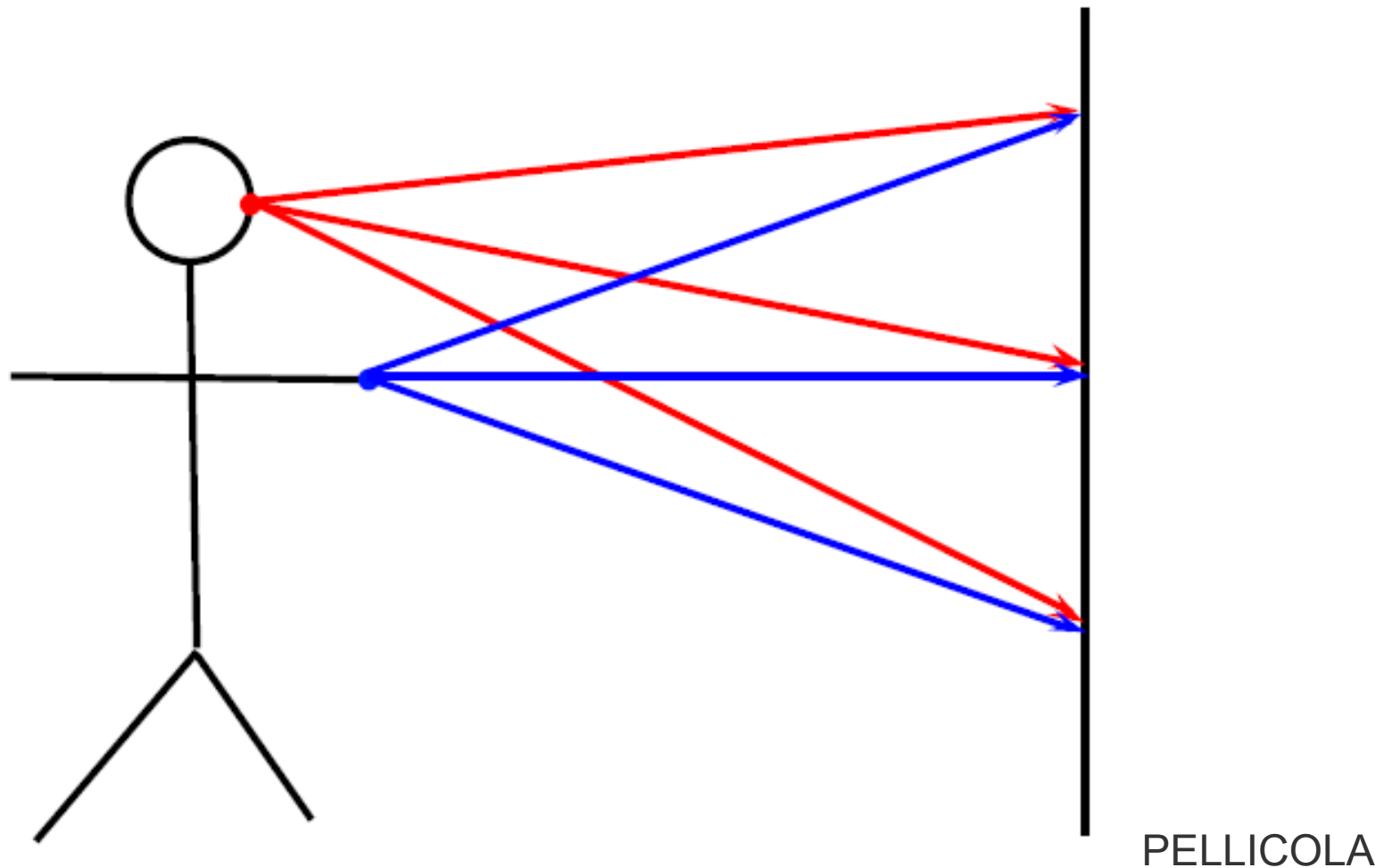
Prima fotografia

- Era noto sin dal medioevo che certi sali di argento si oscurano velocemente sotto la luce del sole.
- Nel 1826 Niepce ottenne la prima vera fotografia esponendo alla luce della carta trattata con cloruro di argento posta all'interno di una camera oscura e fissando poi l'immagine con acido nitrico.



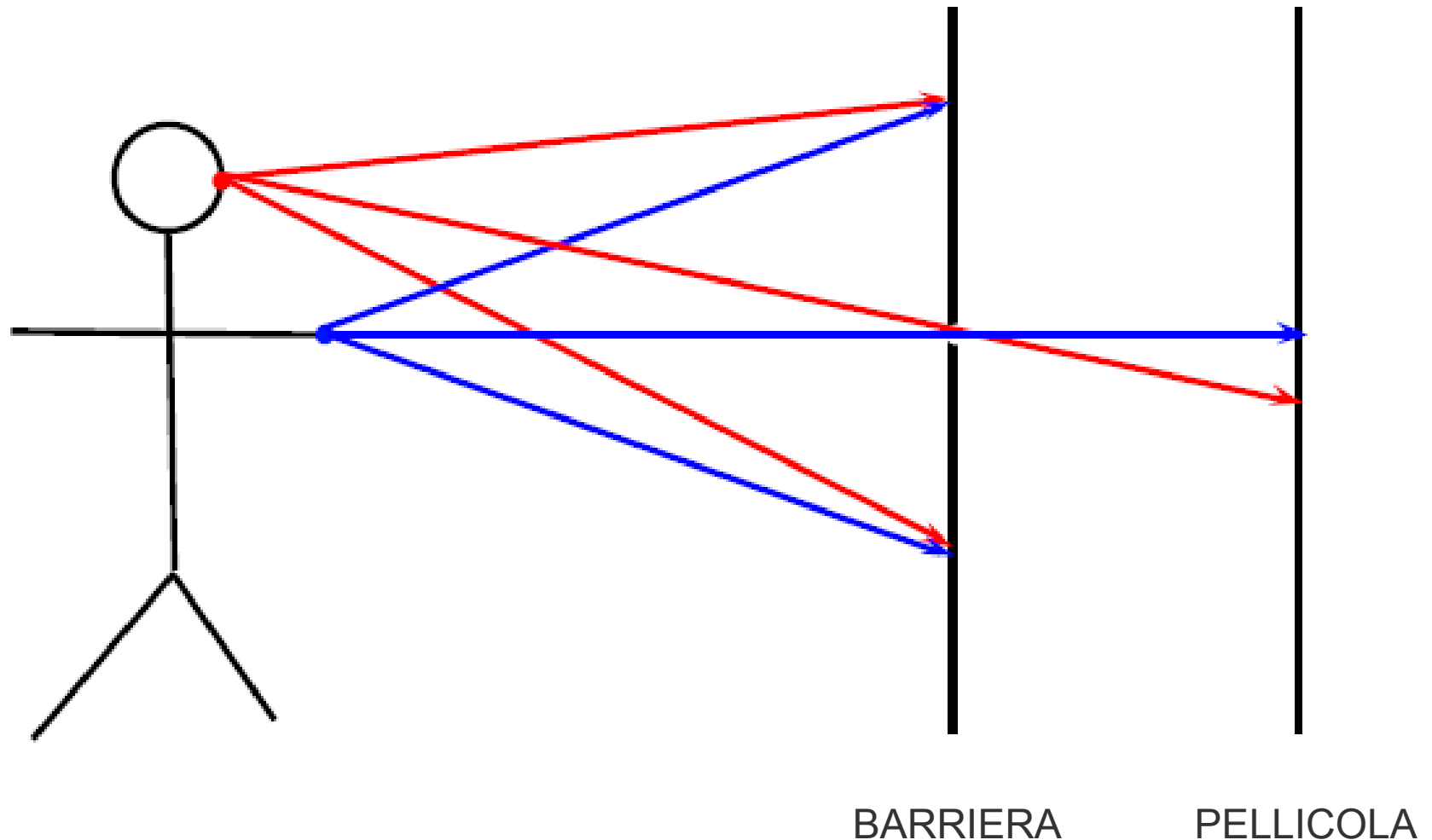
Formazione di immagini

- Che succede ponendo una pellicola di fronte ad un oggetto?



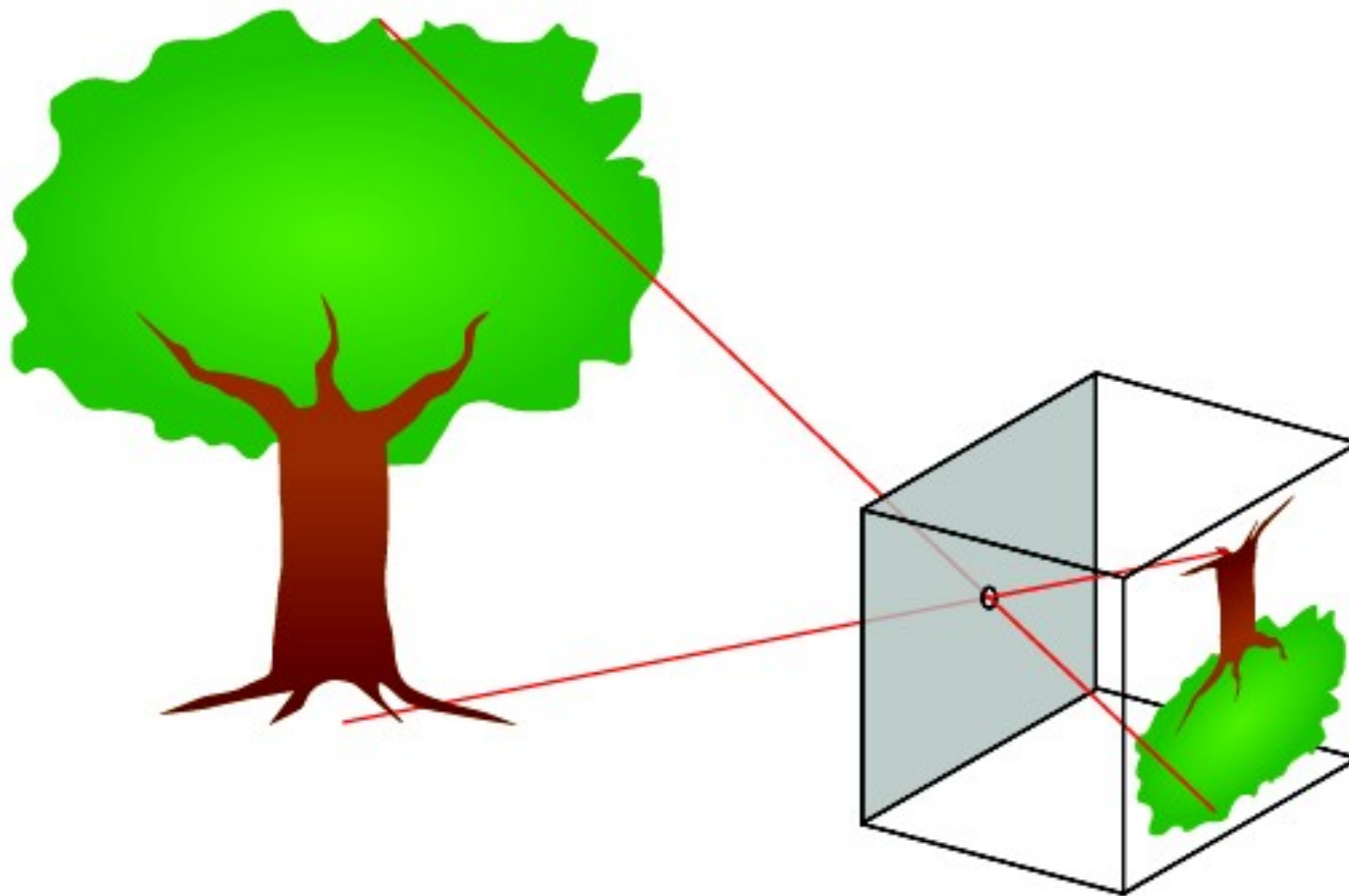
Formazione di immagine

- Aggiungendo una barriera con un foro (foto stenopeico) si riduce la sfocatura.

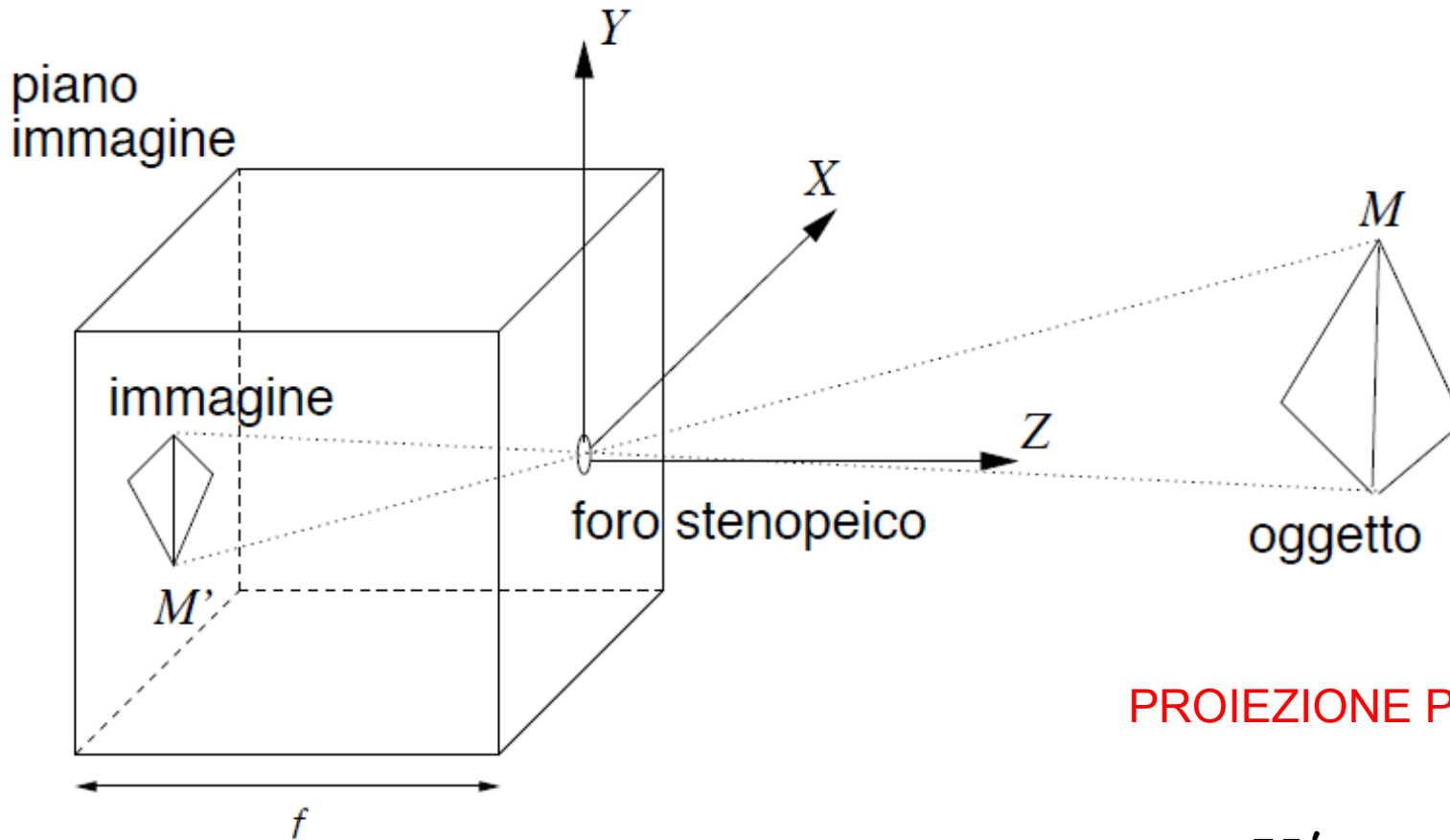


Camera oscura

- Nota ad Aristotele ed Euclide



Modello della camera oscura



PROIEZIONE PROSPETTICA

$$M = (X, Y, Z)$$

$$M' = (X', Y', Z')$$

$$X' = -f \frac{X}{Z}$$

$$Y' = -f \frac{Y}{Z}$$

$$Z' = -f$$

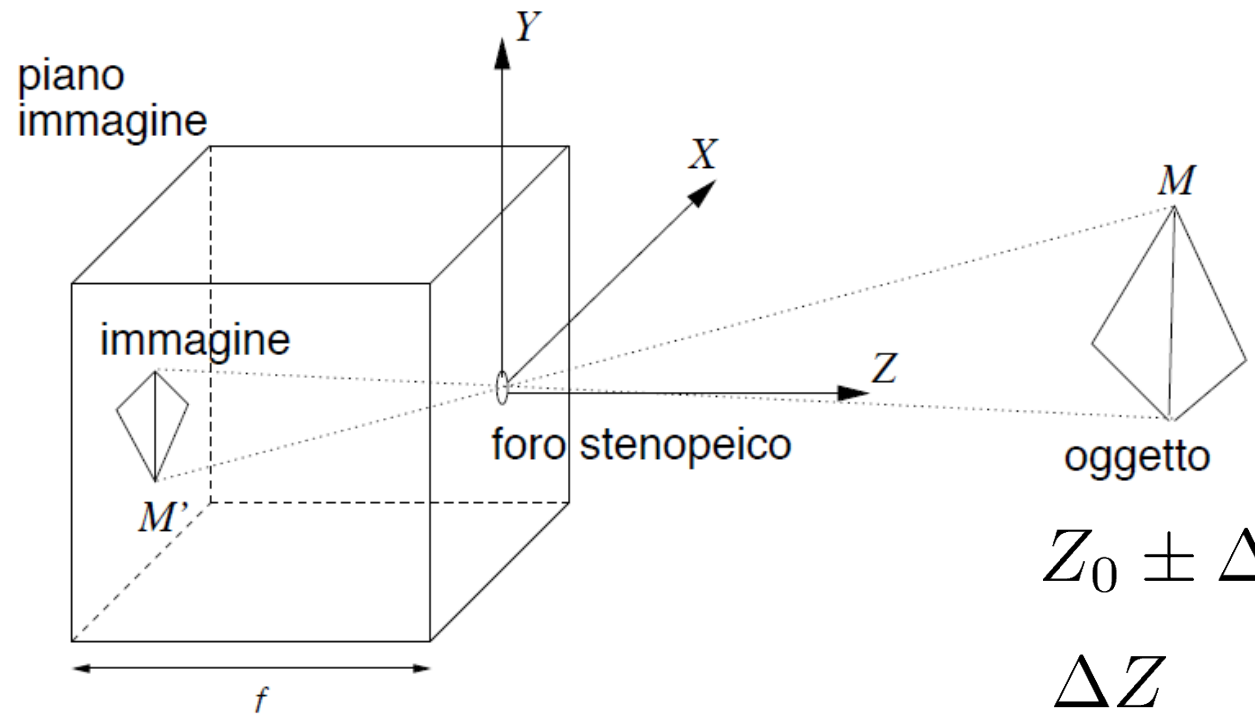
Proiezione ortografica

- Se l'oggetto ha una profondità ininfluente rispetto alla distanza dalla camera possiamo adottare un modello ortografico.

$$X' = -f \frac{X}{Z_0}$$

$$Y' = -f \frac{Y}{Z_0}$$

$$Z' = -f$$

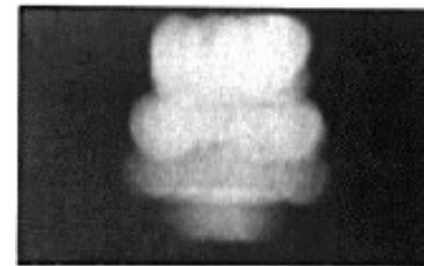
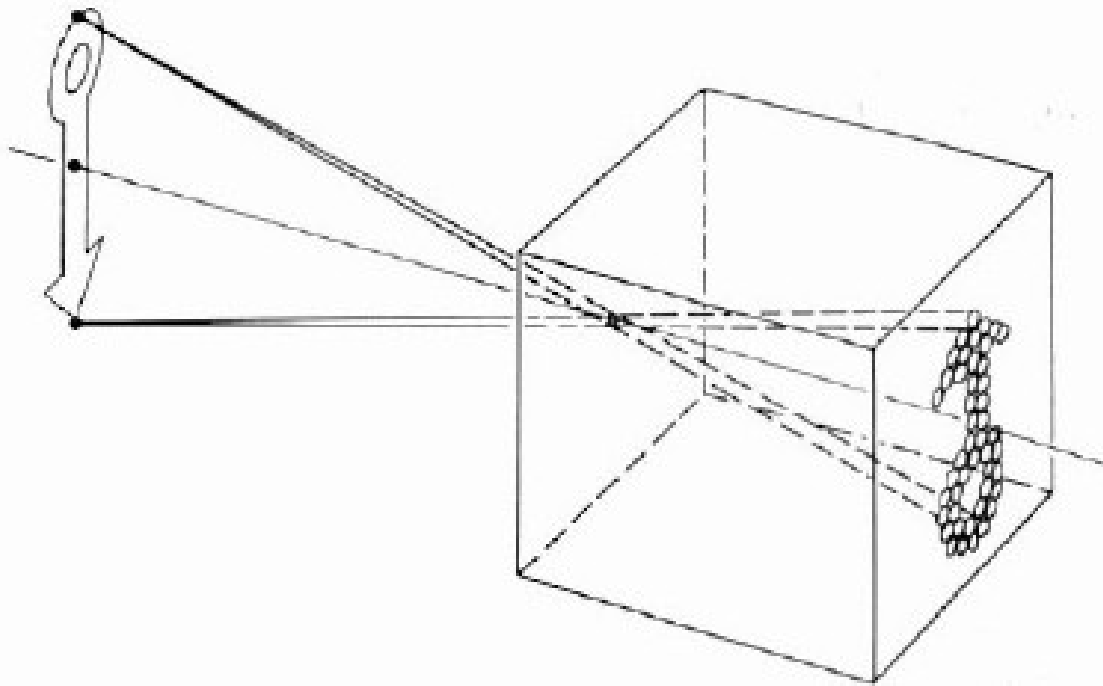


$$Z_0 \pm \Delta Z$$

$$\frac{\Delta Z}{Z_0} \ll 1$$

Problemi della camera oscura

- Un foro troppo grande introduce problemi di sfocatura
- Un foro troppo piccolo fa passare troppa poca luce (tempi di esposizione lunghi)



2 mm



1 mm



0.6mm



0.35 mm



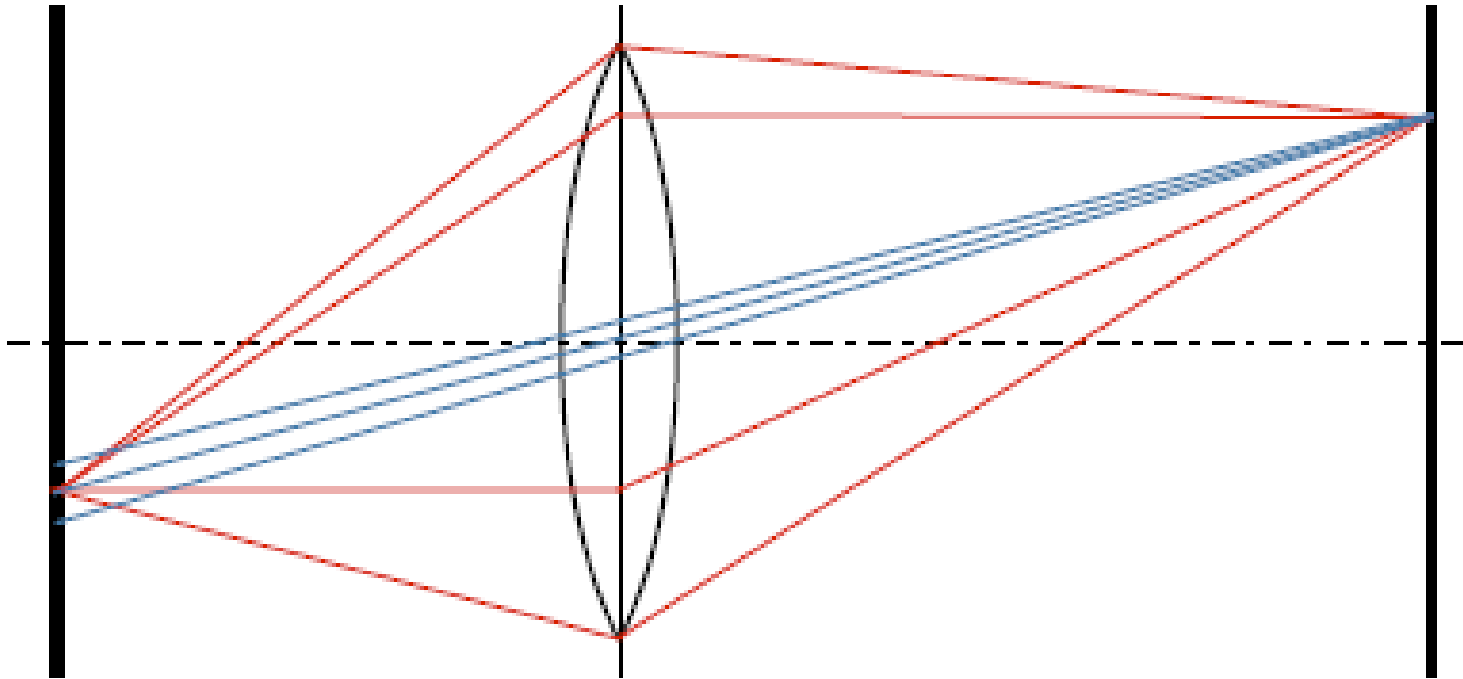
0.15 mm



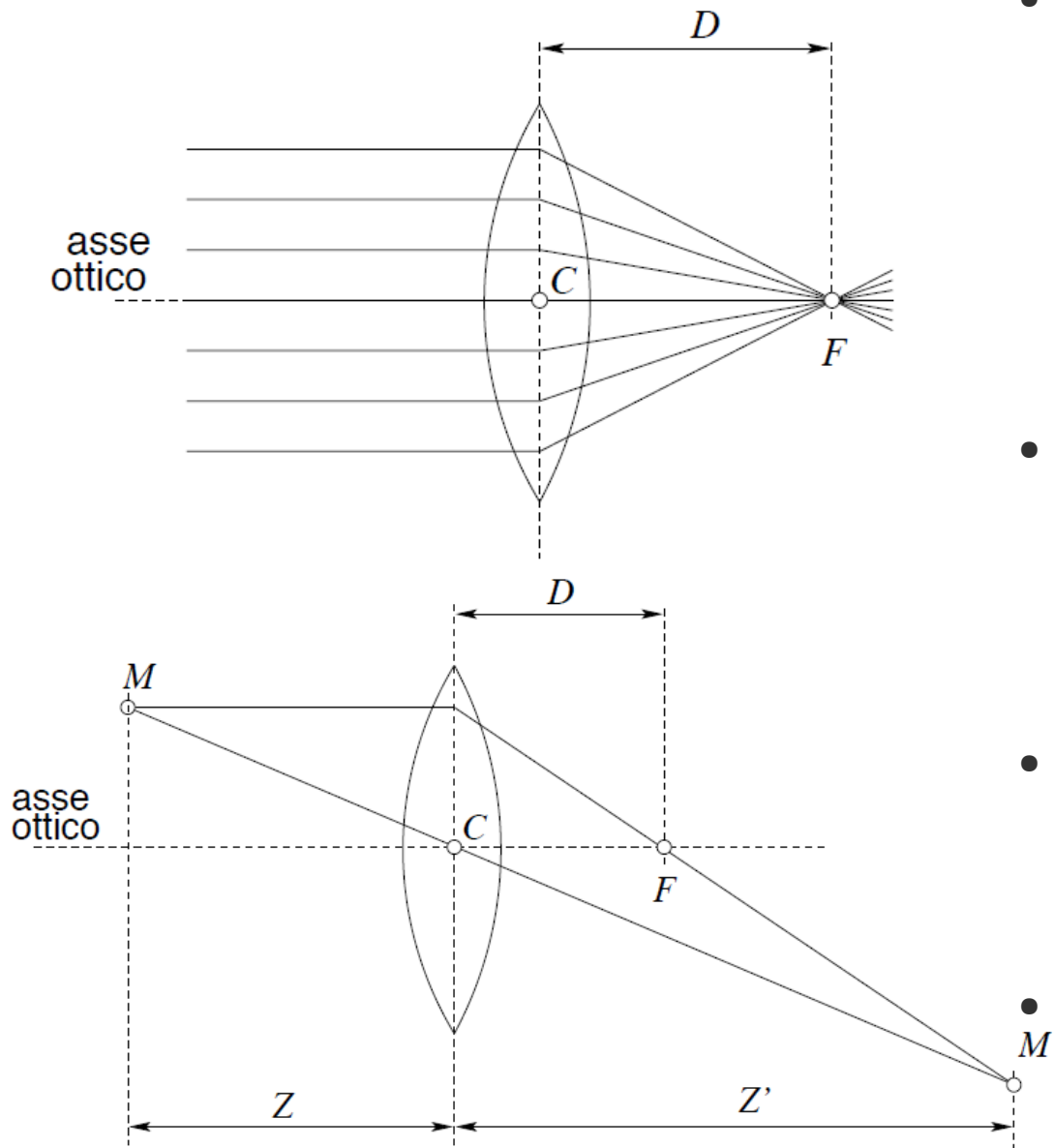
0.07 mm

Lenti

- Le lenti consentono di raccogliere più luce e mantenere la stessa proiezione sui punti a fuoco.
- Per contro, non tutta la scena può essere messa a fuoco allo stesso momento.



Lenti sottili



- Nei punti a fuoco vale la relazione

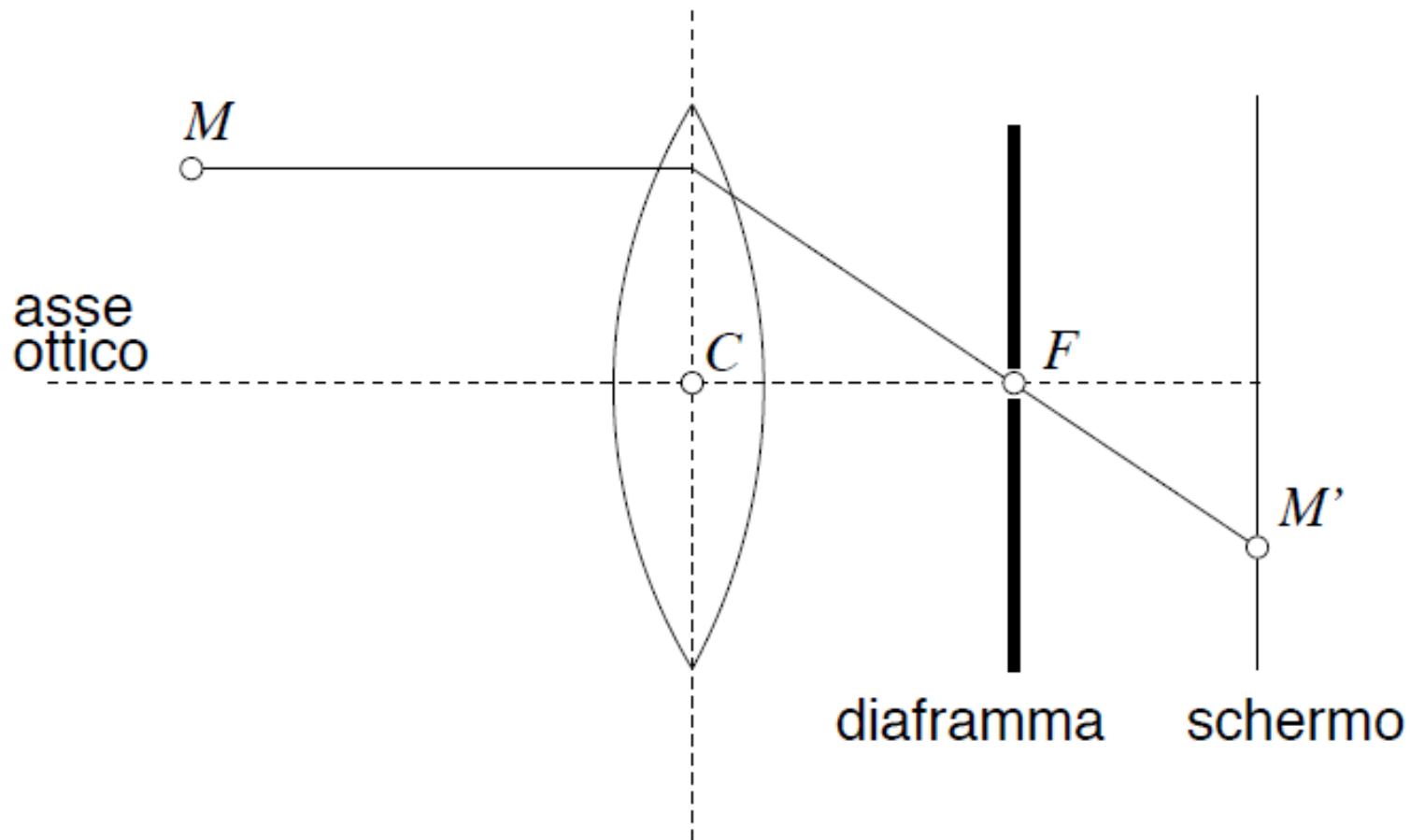
$$\frac{1}{Z} + \frac{1}{Z'} = \frac{1}{D}$$

Nei punti non a fuoco si crea un **cerchio di confusione**.

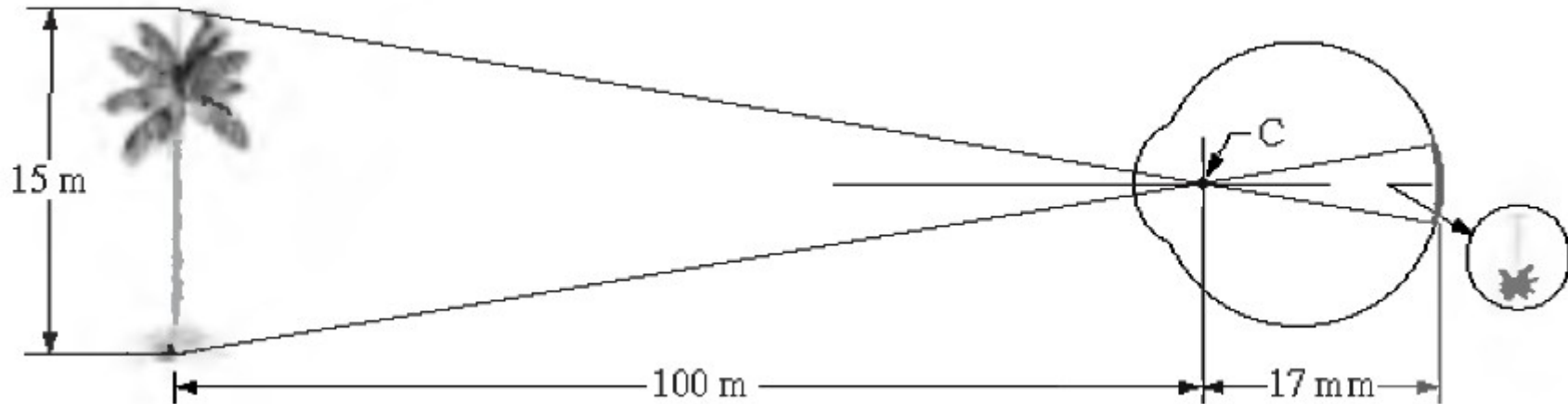
- Finché il cerchio di confusione non supera le dimensioni degli elementi fotosensibili l'immagine è a fuoco.
- L'intervallo di profondità in cui i punti sono a fuoco è detto **profondità di campo**.
- La profondità di campo è inversamente proporzionale al diametro della lente.

Ottica telecentrica

- Con un'ottica telecentrica l'immagine di un punto non dipende dalla profondità.
- So ottiene ponendo una barriera con un foro steopeico in corrispondenza del punto di fuoco.



Formazione dell'immagine nell'occhio



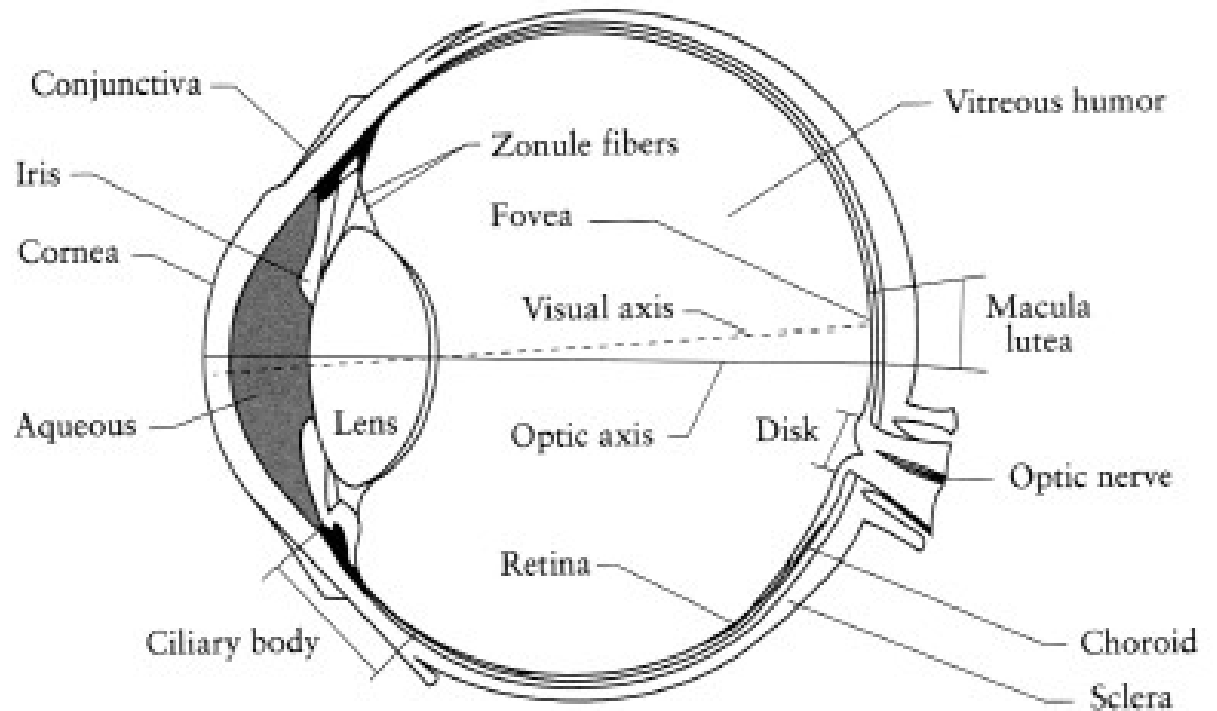
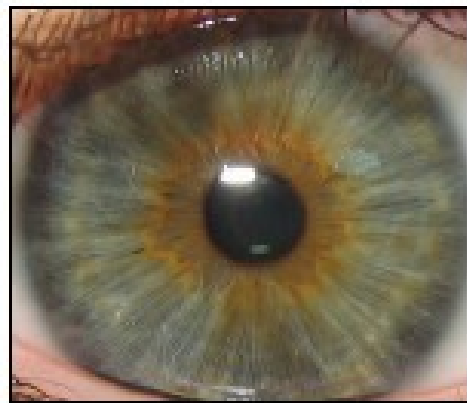
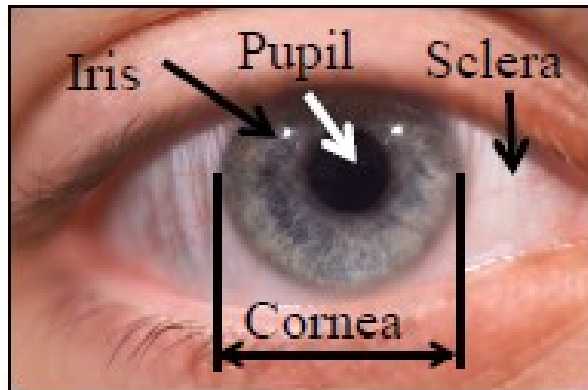
Fotocamera

- Lente ha una distanza focale fissa.
- La messa a fuoco avviene variando la distanza tra lente e piano immagine, in cui la pellicola/chip è posizionata.

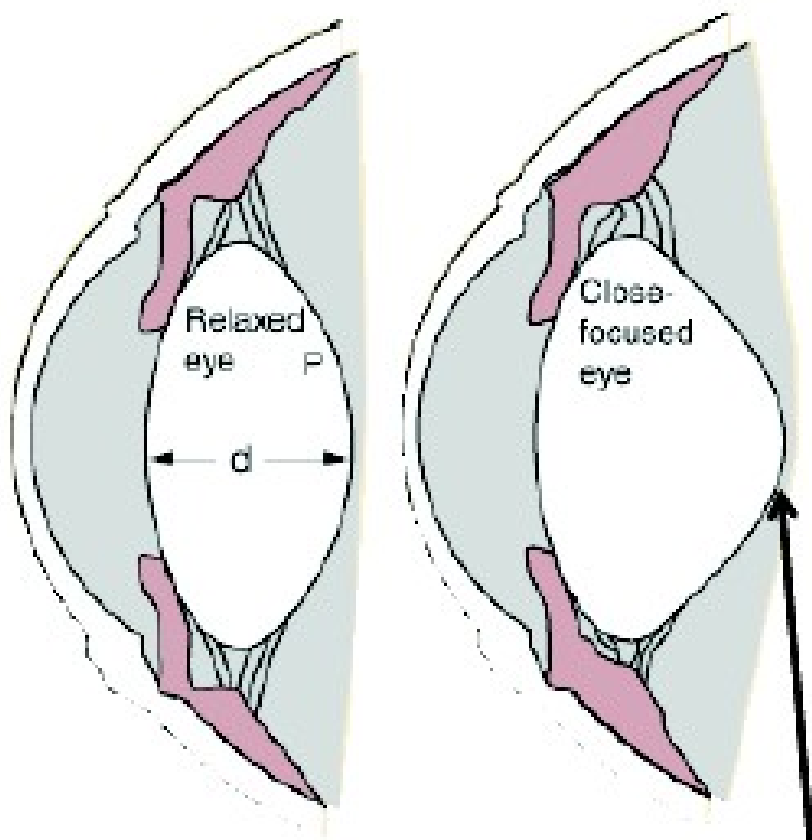
Occhio

- Distanza tra lente e regione immagine (retina) è fissa.
- La forma della lente viene variata per ottenere una messa a fuoco dell'immagine.

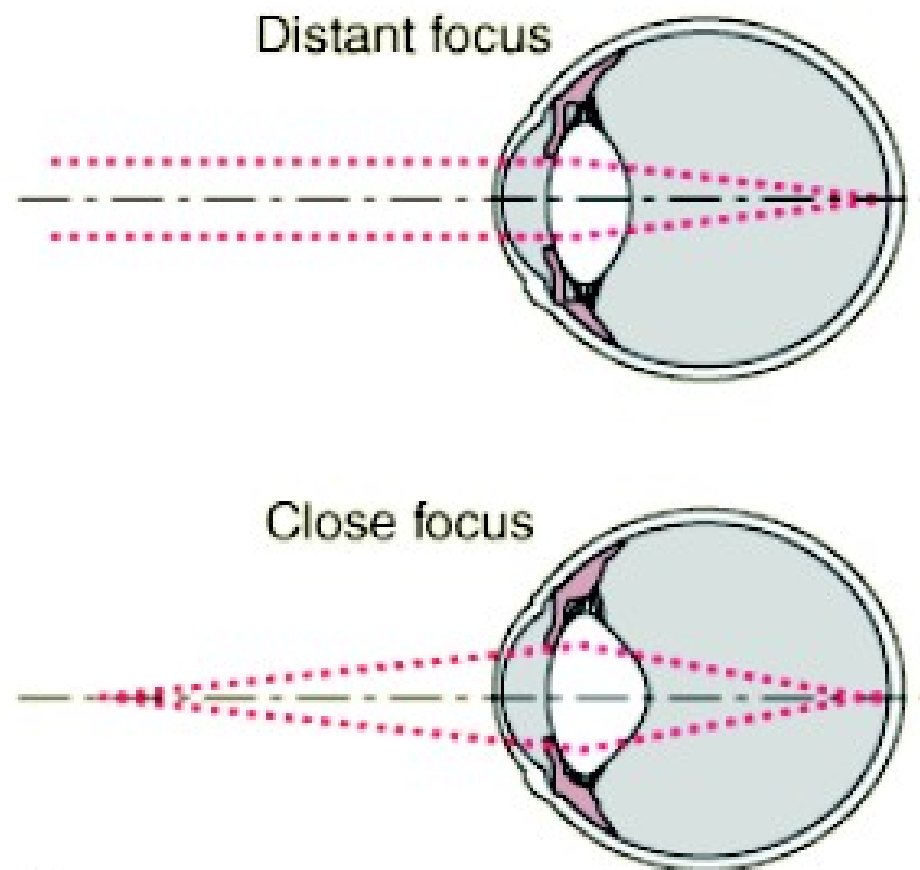
Formazione dell'immagine nell'occhio



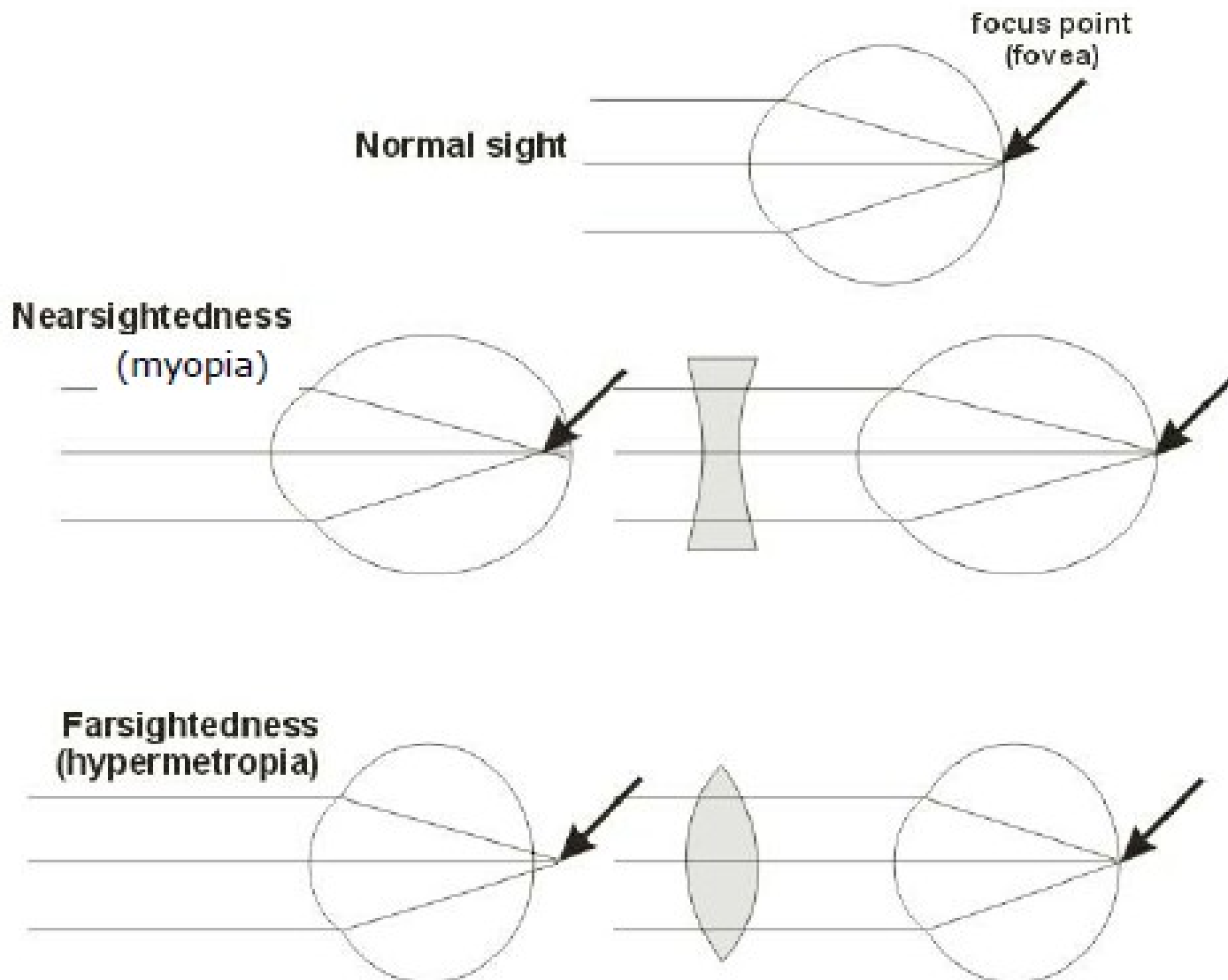
Formazione dell'immagine nell'occhio



shorter focal length



Miopia e ipermetropia



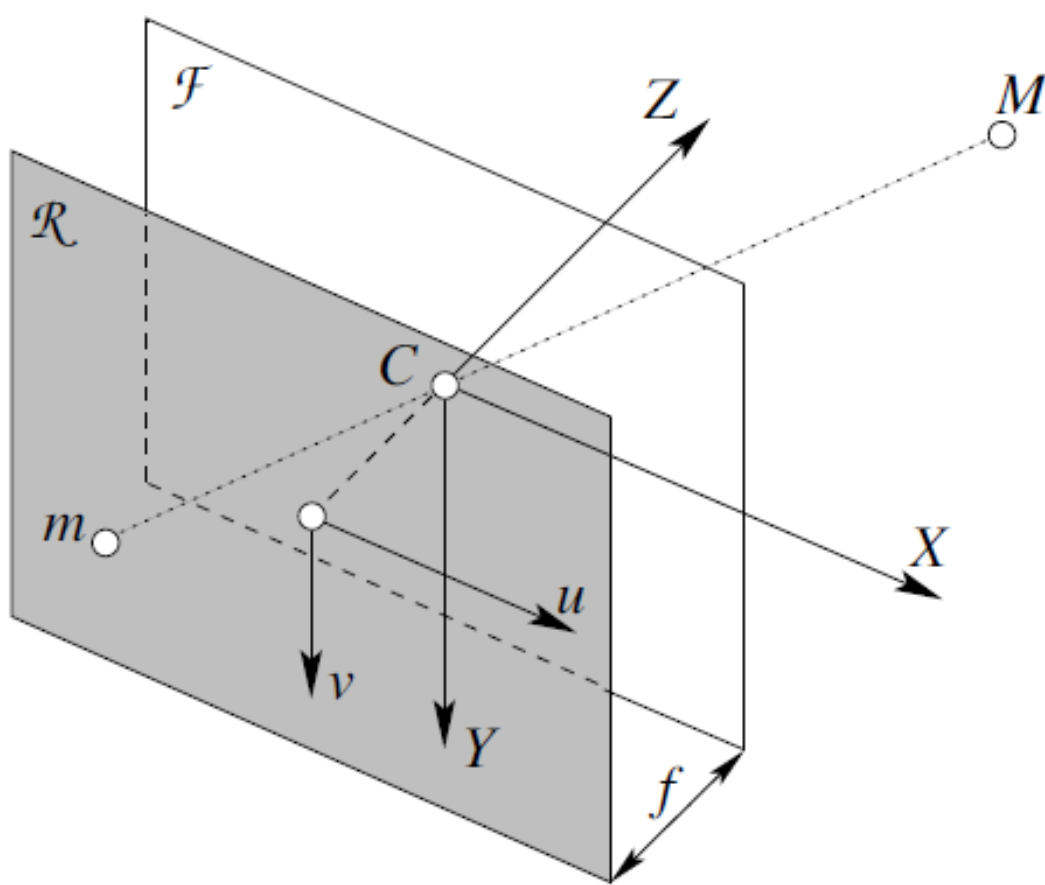
Immagini digitali

- Una fotocamera digitale è composta dall'ottica, approssimata con una lente sottile, e una matrice CCD (Charge-Coupled Device) o CMOS che costituisce il piano immagine.
- La matrice del CCD (o CMOS) viene convertita in immagine, ovvero una matrice $M \times N$ di valori interi.
- La dimensione $n \times m$ della matrice CCD non è necessariamente la stessa dell'immagine $N \times M$. Ad ogni pixel corrisponde un'area rettangolare del CCD (detta **impronta del pixel**).

$$u_{\text{pix}} = \frac{N}{n} u_{\text{CCD}}$$

$$u_{\text{pix}} = \frac{M}{m} u_{\text{CCD}}$$

Modello della fotocamera



$$\mathbf{m} = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \quad \mathbf{M} = \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

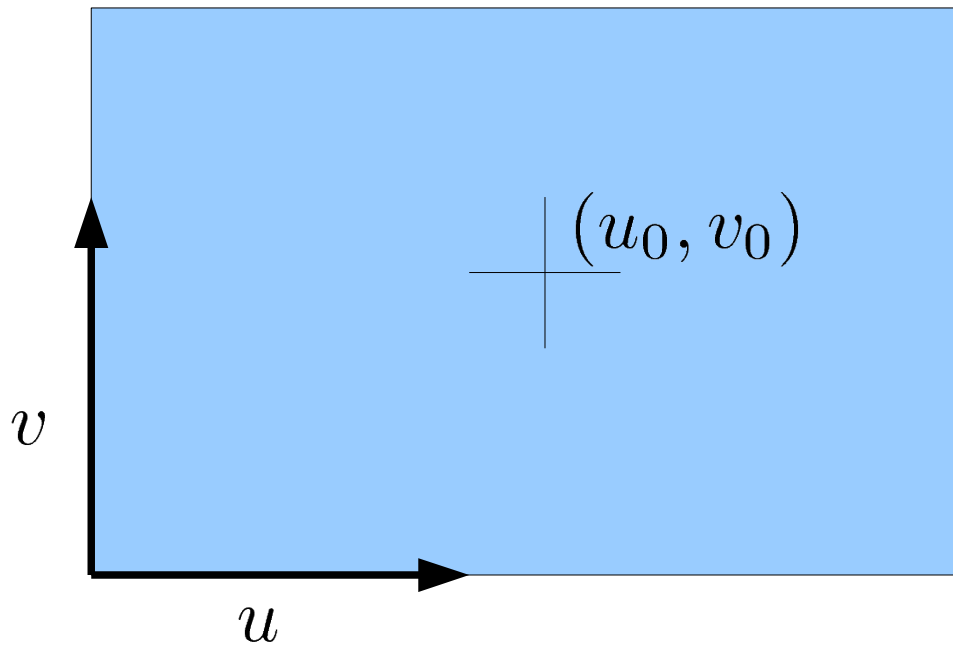
$$\begin{cases} u = -f \frac{x}{z} \\ v = -f \frac{y}{z} \end{cases}$$

$$z\mathbf{m} = K[I|\mathbf{0}]\mathbf{M}$$

$$\mathbf{m} \simeq K[I|\mathbf{0}]\mathbf{M}$$

$$K = \begin{bmatrix} -f & 0 & 0 \\ 0 & -f & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Pixelizzazione



$$\begin{cases} u = -f k_u \frac{x}{z} + u_0 \\ v = -f k_v \frac{y}{z} + v_0 \end{cases}$$

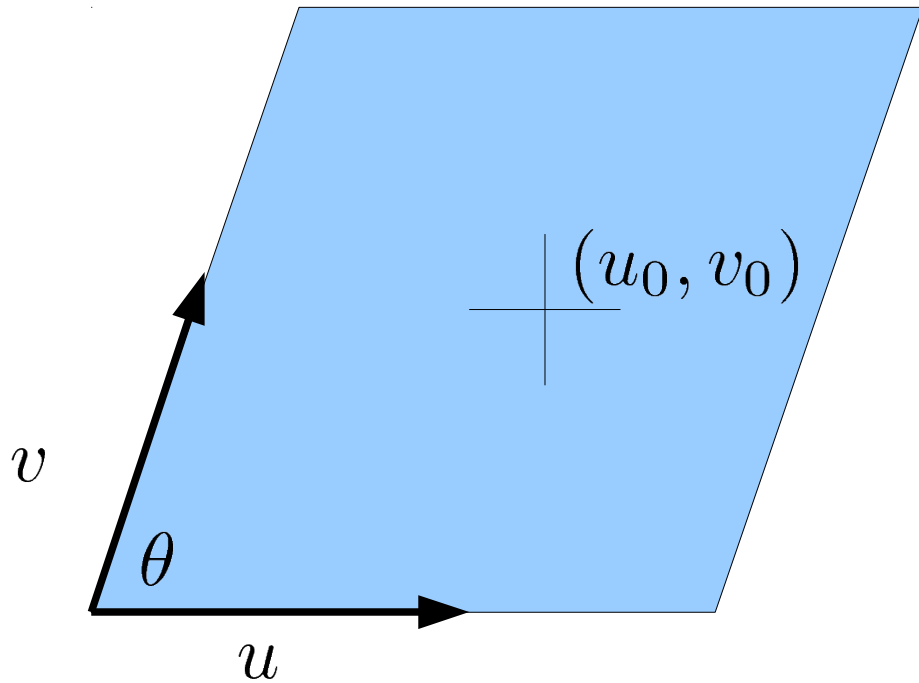
$$\alpha_u = -f k_u$$

$$\alpha_v = -f k_v$$

$$\mathbf{m} \simeq K [I | \mathbf{0}] \mathbf{M}$$

$$K = \begin{bmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

Skewness



$$\mathbf{m} \simeq K [I | \mathbf{0}] \mathbf{M}$$

$$K = \begin{bmatrix} \alpha_u & -\alpha_u / \tan \theta & u_0 \\ 0 & \alpha_v / \sin \theta & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

Parametri intrinseci

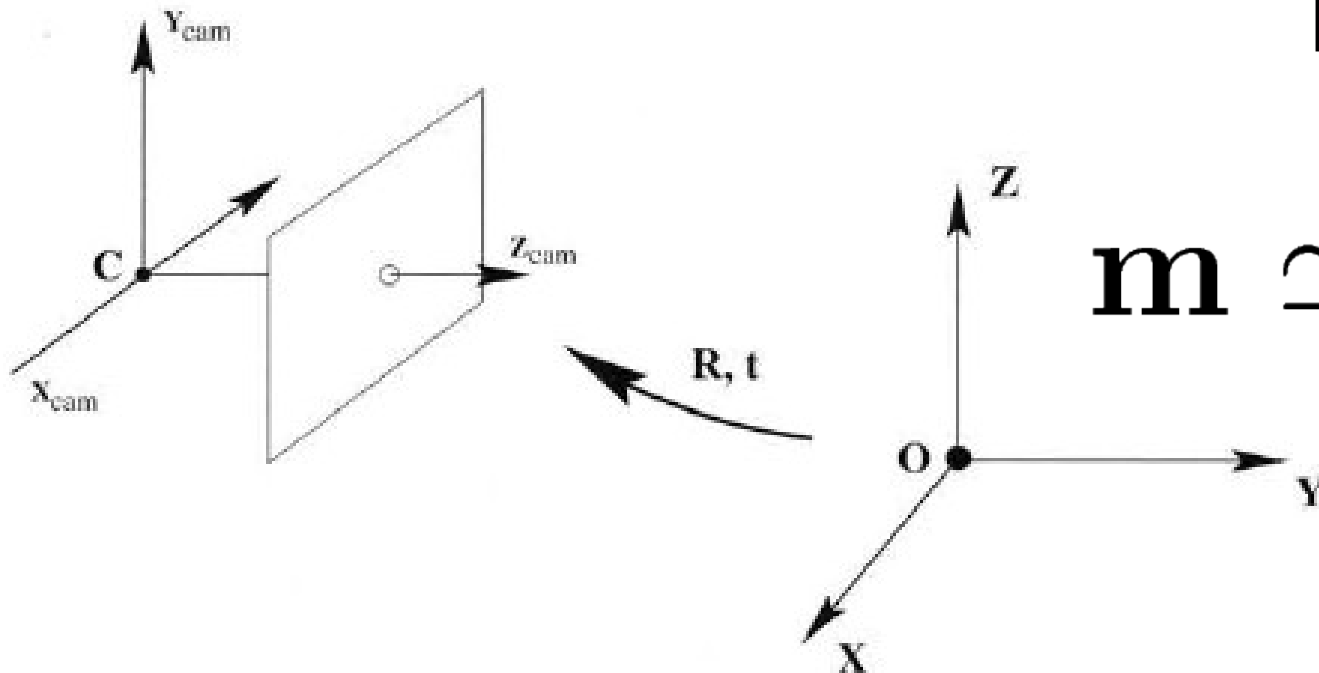
- I parametri intrinseci sono i parametri che sono codificati nella matrice K .
- Rappresentano parametri che caratterizzano la singola fotocamera, indipendentemente dal mondo che la circonda.

$$K = \begin{bmatrix} \alpha_u & -\alpha_u / \tan \theta & u_0 \\ 0 & \alpha_v / \sin \theta & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

Parametri estrinseci

- I parametri estrinseci caratterizzano la trasformazione tra coordinate camera e coordinate mondo.

$$\mathbf{M}_c = \mathbf{GM} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0}^\top & 1 \end{bmatrix} \mathbf{M}$$



$$\mathbf{m} \simeq \mathbf{K} [\mathbf{R} | \mathbf{t}] \mathbf{M}$$

Matrice di rotazione

$$R = R_{z,\phi} R_{y,\theta} R_{x,\psi}$$

$$R_{x,\psi} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \psi & -\sin \psi \\ 0 & \sin \psi & \cos \psi \end{bmatrix}$$

$$R_{y,\theta} = \begin{bmatrix} \cos \theta & 0 & \sin \theta \\ 0 & 1 & 0 \\ -\sin \theta & 0 & \cos \theta \end{bmatrix}$$

$$R_{z,\phi} = \begin{bmatrix} \cos \phi & -\sin \phi & 0 \\ \sin \phi & \cos \phi & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Matrice di rotazione rispetto ad un asse

$$R = I + \sin \theta [\mathbf{u}]_{\times} + (1 - \cos \theta) [\mathbf{u}]_{\times}^2$$

$$[\mathbf{u}]_{\times} = \begin{bmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{bmatrix}$$

$$\|\mathbf{u}\| = 1$$

\mathbf{u} è il versore unitario che individua l'asse di rotazione (che passa per l'origine).

Matrice di proiezione prospettica (MPP)

$$P = K[R|\mathbf{t}]$$

$$\mathbf{m} \simeq \begin{bmatrix} \mathbf{p}_1^\top \mathbf{M} \\ \mathbf{p}_2^\top \mathbf{M} \\ \mathbf{p}_3^\top \mathbf{M} \end{bmatrix} = \begin{bmatrix} \mathbf{p}_1^\top \\ \mathbf{p}_2^\top \\ \mathbf{p}_3^\top \end{bmatrix} \mathbf{M} = P\mathbf{M}$$

$$\begin{cases} u = \frac{\mathbf{p}_1^\top \mathbf{M}}{\mathbf{p}_3^\top \mathbf{M}} \\ v = \frac{\mathbf{p}_2^\top \mathbf{M}}{\mathbf{p}_3^\top \mathbf{M}} \end{cases}$$

Centro ottico

- Il piano focale contiene i punti che si proiettano all'infinito ed ha quindi equazione:

$$\mathbf{p}_3^\top \mathbf{M} = 0$$

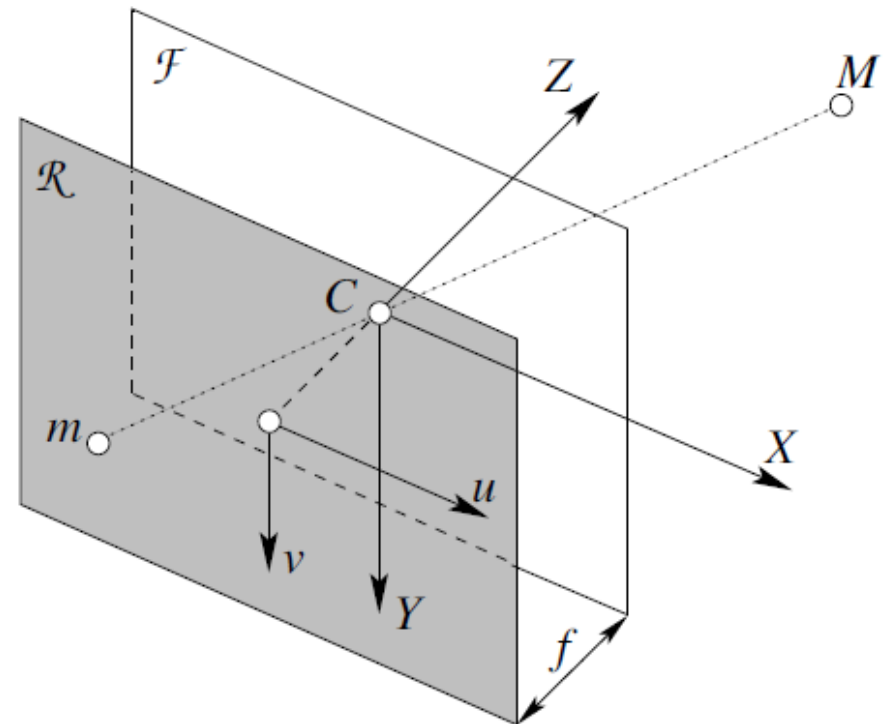
- I piani di che si proiettano nell'immagine sugli assi $u=0$ e $v=0$ sono dati da

$$\mathbf{p}_1^\top \mathbf{M} = 0$$

$$\mathbf{p}_2^\top \mathbf{M} = 0$$

- Il centro ottico \mathbf{C} è dato dall'intersezione di questi 3 piani.

$$PC = \mathbf{0}$$



Centro ottico

- Riscriviamo la MPP nel seguente modo

$$P = [Q | \mathbf{q}]$$

e ricordando che

$$\mathbf{C} = \begin{bmatrix} \tilde{\mathbf{C}} \\ 1 \end{bmatrix}$$

otteniamo

$$\mathbf{0} = Q\tilde{\mathbf{C}} + \mathbf{q}$$

$$\tilde{\mathbf{C}} = -Q^{-1}\mathbf{q}$$

Raggio ottico

- Il raggio ottico del punto \mathbf{m} è la linea retta che passa per il centro ottico \mathbf{C} ed \mathbf{m} stesso.

$$\{\mathbf{M} : \mathbf{m} \simeq P\mathbf{M}\}$$

- Oltre al centro ottico, un altro punto (ideale) del raggio è

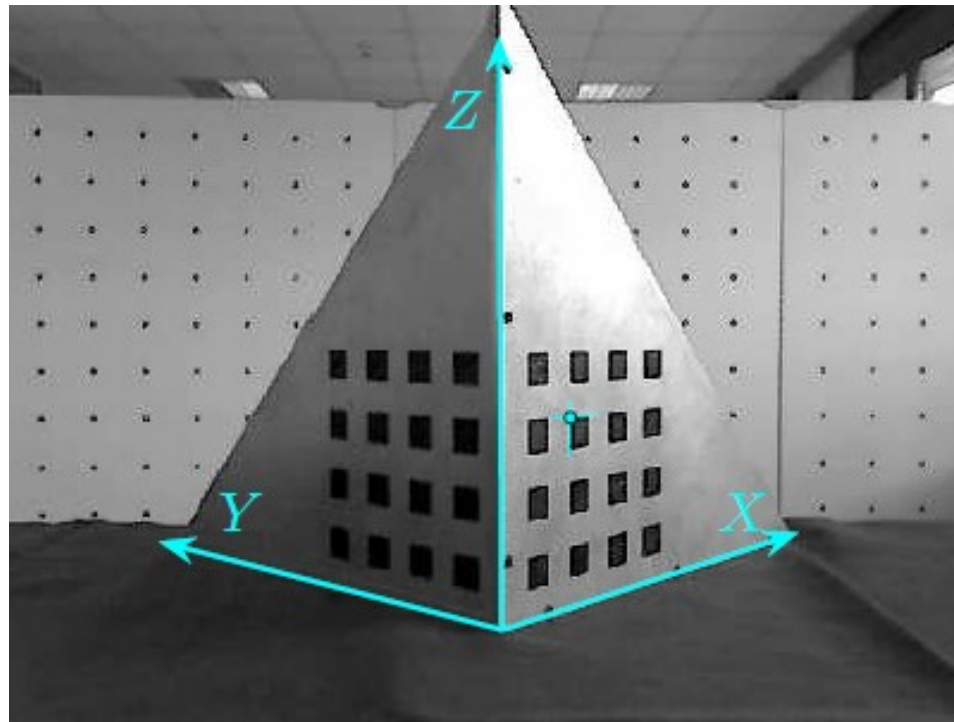
$$\begin{bmatrix} Q^{-1}\mathbf{m} \\ 0 \end{bmatrix}$$

- L'equazione parametrica del raggio ottico è la seguente:

$$\mathbf{M} = \mathbf{C} + \lambda \begin{bmatrix} Q^{-1}\mathbf{m} \\ 0 \end{bmatrix}, \quad \lambda \in \mathbb{R} \cup \{\infty\}$$

Calibrazione

- La calibrazione consiste nel misurare i parametri intrinseci ed estrinseci della camera.
- L'idea è che conoscendo le proiezioni di punti 3D di coordinate note, sia possibile ottenere i parametri incogniti risolvendo le equazioni della proiezione prospettica.



Metodo DLT

- Dati n punti di calibrazione non coplanari, ciascuna corrispondenza tra un punto dell'immagine $\mathbf{m}_i = [u_i, v_i, 1]$ e uno della scena \mathbf{M}_i fornisce una coppia di equazioni:

$$\begin{cases} \mathbf{p}_1^\top \mathbf{M}_i - u_i \mathbf{p}_3^\top \mathbf{M}_i = 0 \\ \mathbf{p}_2^\top \mathbf{M}_i - v_i \mathbf{p}_3^\top \mathbf{M}_i = 0 \end{cases}$$

che in forma matriciale diviene

$$\begin{bmatrix} \mathbf{M}_i^\top & \mathbf{0} & -u_i \mathbf{M}_i^\top \\ \mathbf{0} & -\mathbf{M}_i^\top & v_i \mathbf{M}_i^\top \end{bmatrix} \begin{bmatrix} \mathbf{p}_1 \\ \mathbf{p}_2 \\ \mathbf{p}_3 \end{bmatrix} = \mathbf{0}_{2 \times 1}$$

- Da risolvere ai minimi quadrati con almeno 6 punti.

Metodo non lineare

- Il metodo DLT è lineare, veloce, ma minimizza un errore algebrico $\|A\mathbf{x}\|$ che non ha significato geometrico.
- Una precisione maggiore può essere ottenuta minimizzando un errore geometrico.
- Per esempio possiamo minimizzare l'errore di proiezione.

$$\mathcal{E}(P) = \sum_{i=1}^n \left(\frac{\mathbf{p}_1^\top \mathbf{M}_i}{\mathbf{p}_3^\top \mathbf{M}_i} - u_i \right)^2 + \left(\frac{\mathbf{p}_2^\top \mathbf{M}_i}{\mathbf{p}_3^\top \mathbf{M}_i} - v_i \right)^2$$

Estrazione dei parametri

- Data una matrice $P=[Q|\mathbf{q}]$ di rango 3 possiamo decomporla come

$$P \simeq K[R|\mathbf{t}]$$

- Notiamo che $Q=KR$ dove K è triangolare superiore ed R ortogonale. Se invertiamo Q otteniamo

$$Q^{-1} = US = R^{-1}K^{-1}$$

che altro non è che la decomposizione QR, con U ortogonale e S triangolare superiore

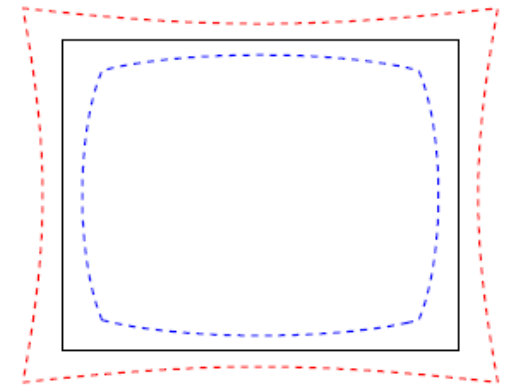
- Infine ricaviamo $\mathbf{t} = K^{-1}\mathbf{q} = S\mathbf{q}$

Distorsione radiale

- Un modello più accurato della camera deve tenere conto della distorsione radiale delle lenti.
- Il modello standard è una trasformazione delle coordinate corrette (u,v) alle coordinate reali distorte

$$\begin{cases} \hat{u} = (u - u_0)(1 + k_1 r_d^2) + u_0 \\ \hat{v} = (v - v_0)(1 + k_1 r_d^2) + v_0 \end{cases}$$

$$r_d^2 = \left(\frac{u - u_0}{\alpha_u} \right)^2 + \left(\frac{v - v_0}{\alpha_v} \right)^2$$



centro immagine

$$(u_0, v_0)$$