LABORATORIO DI ROBOTICA E MECCATRONICA M

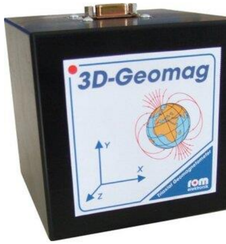
Docenti: Marco Carricato, Edoardo Idà

Tutore: Simone Comari

A.A. 2021-22

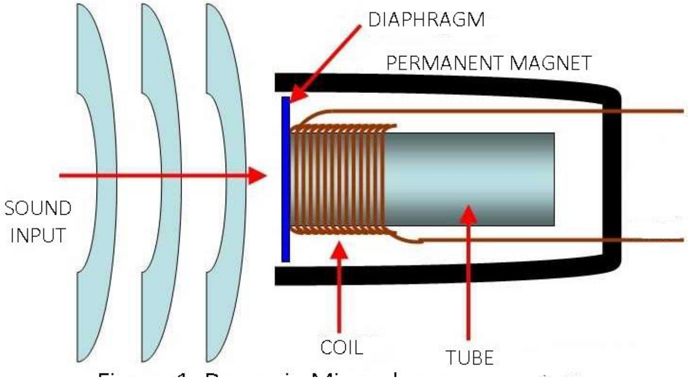
Tipologie di sensori

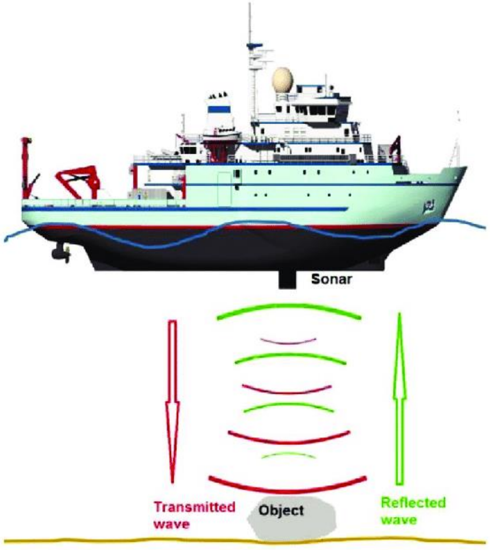
- Sensori **propriocettivi**
  - Misurano valori interni al sistema (robot)
  - E.g. *velocità/corrente dei motori, velocità delle ruote (rover), orientamento del robot, stato della batteria, etc..*
- Sensori **esterocettivi**
  - Estraggono informazioni dall'ambiente circostante il robot
  - E.g. *distanza da oggetti, intensità della luce ambientale, rumore, etc..*




Tipologie di sensori

- Sensori **passivi**
  - Misurano l’energia proveniente direttamente dall’ambiente
  - E.g. *fotocamere, microfoni, etc..*





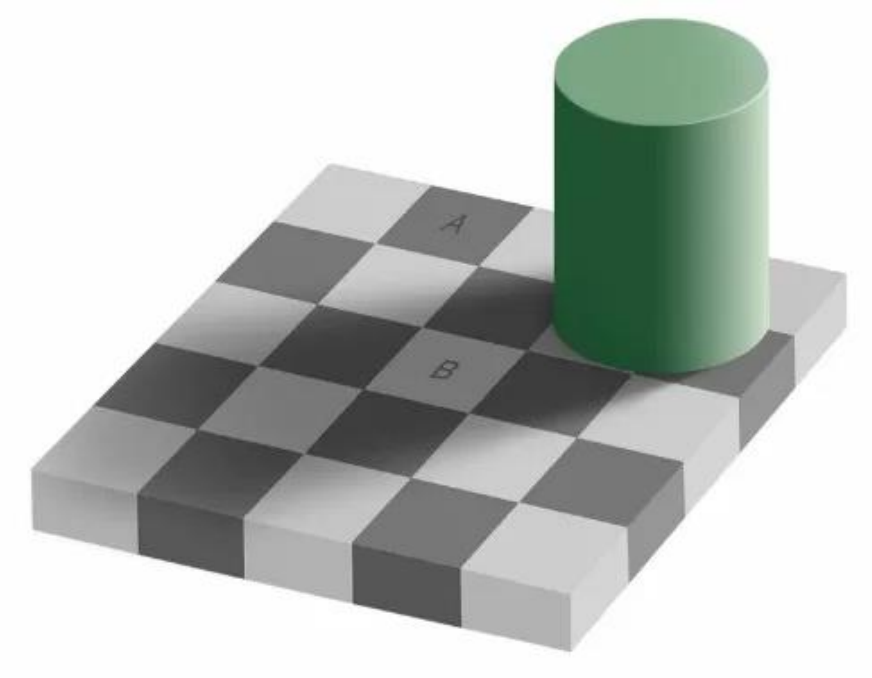
- *Sensori attivi*
  - Emettono energia e misurano la reazione dell’ambiente
  - E.g. *sonar, radar, lidar, etc..*

Sensori di visione

Dalla vista alla visione artificiale

- La **vista** è il nostro senso più importante per percepire e navigare il mondo che ci circonda
- La retina, in circa 10 cm^2, contiene milioni di **fotorecettori** e immagazzina dati ad un impressionante frequenza di circa **~3 Gbytes/s**
- Una grossa porzione del nostro cervello è dedicata a processare i segnali provenienti dai nostri occhi (circa 60 miliardi di neuroni)
- Il nostro sistema di visione è altamente sofisticato
  - Può interpretare immagini sotto una vasta gamma di condizioni, anche in presenza di scarsi indizi
  - Riesce a processare in tempi rapidissimi l'enorme quantità di informazioni presenti in ogni singola immagine
  - Ma non sempre è infallibile..

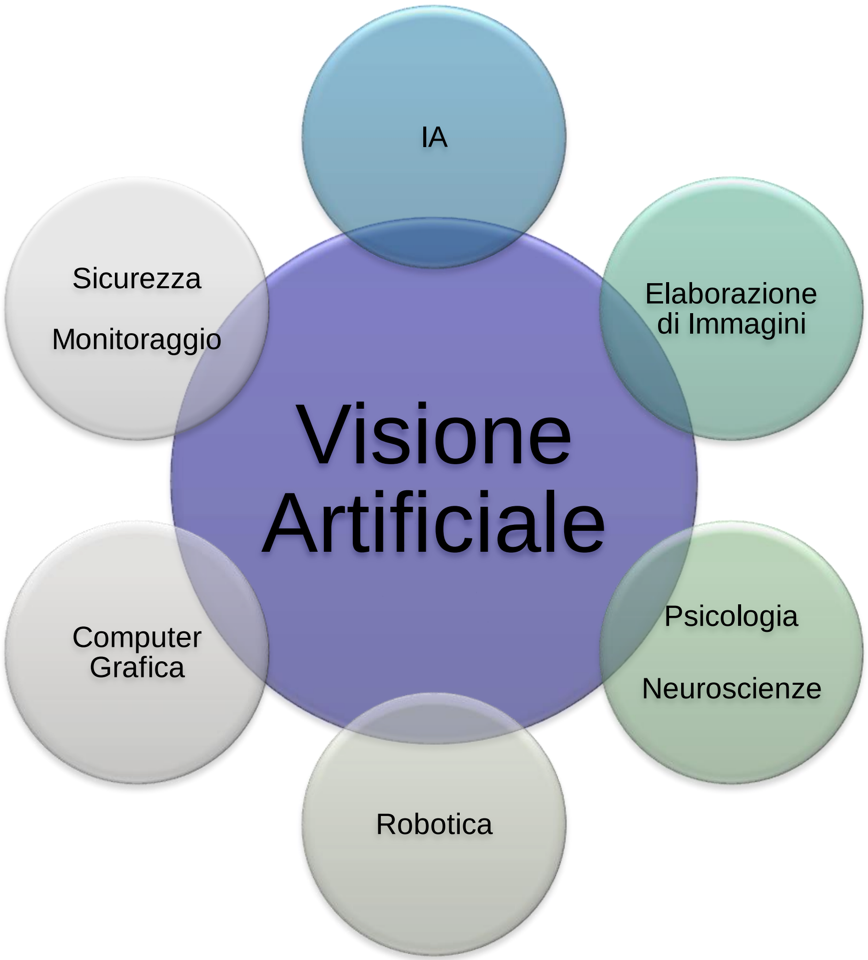
Dalla vista alla visione artificiale

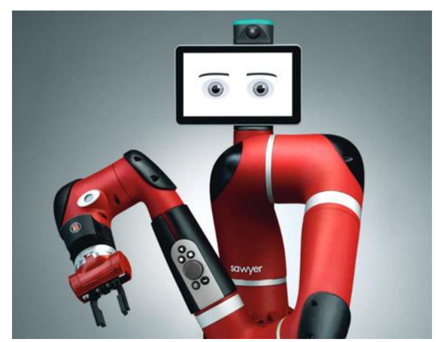


Sensori di visione: fotocamere

- Sensore esterocettivo
- Sensore passivo
- Vantaggi:
  - Altamente descrittive (contengono molte informazioni)
  - Compattezza
  - Compatibilità
  - Basso costo
  - Diffusione
  - Disponibilità
  - HW avanzato per supportare l’elaborazione di immagini

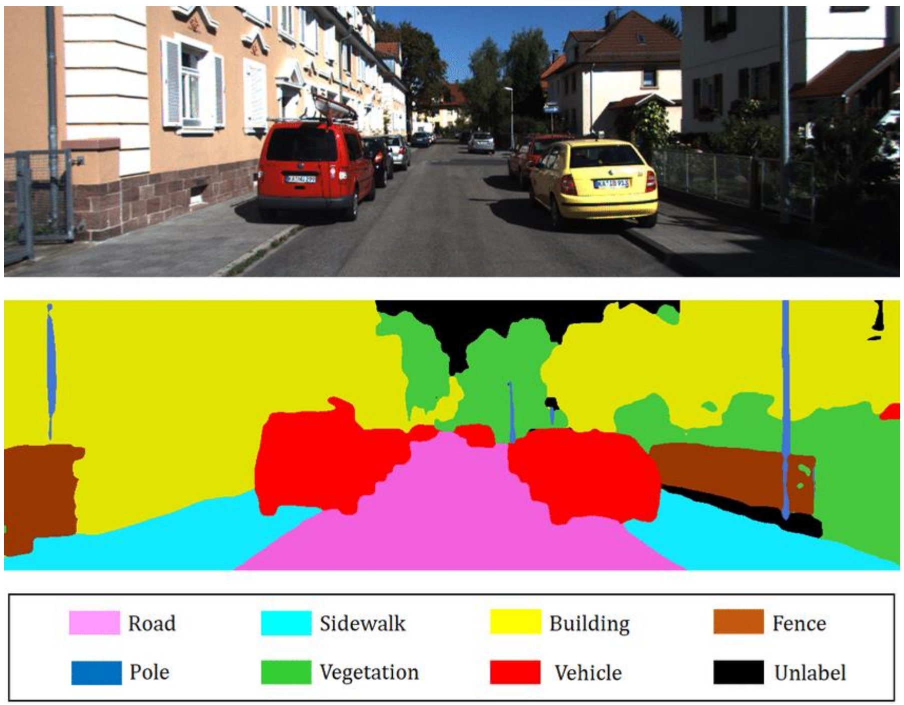
Sensori di visione: applicazioni



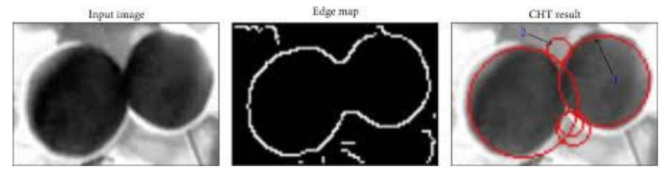


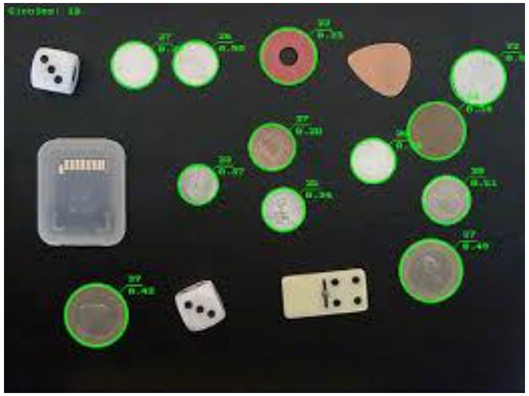
Visione artificiale: cos'è?

- Estrazione automatica di informazioni "significative" da immagini e video
  - varia a seconda dell'applicazione



Semantica

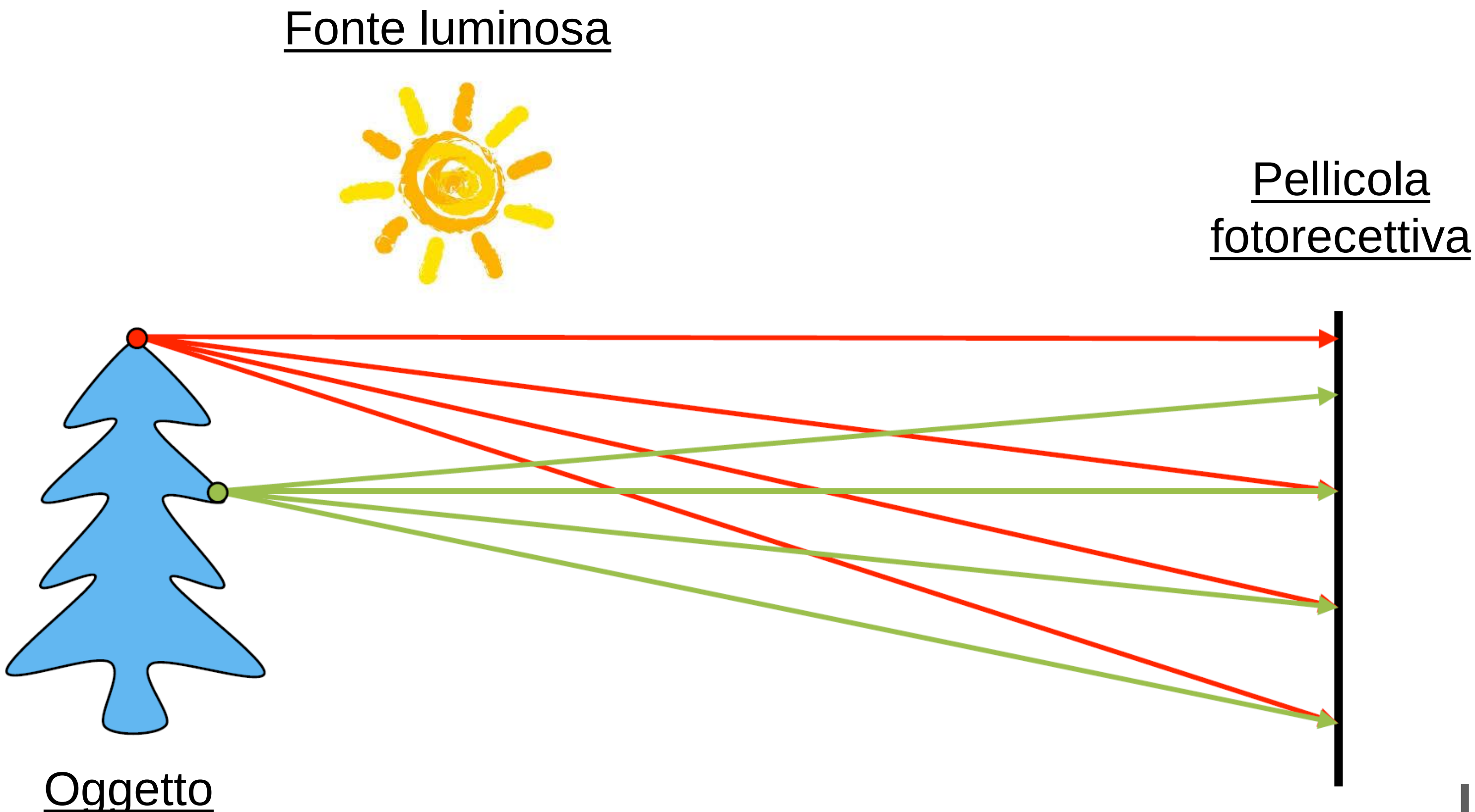




Geometria

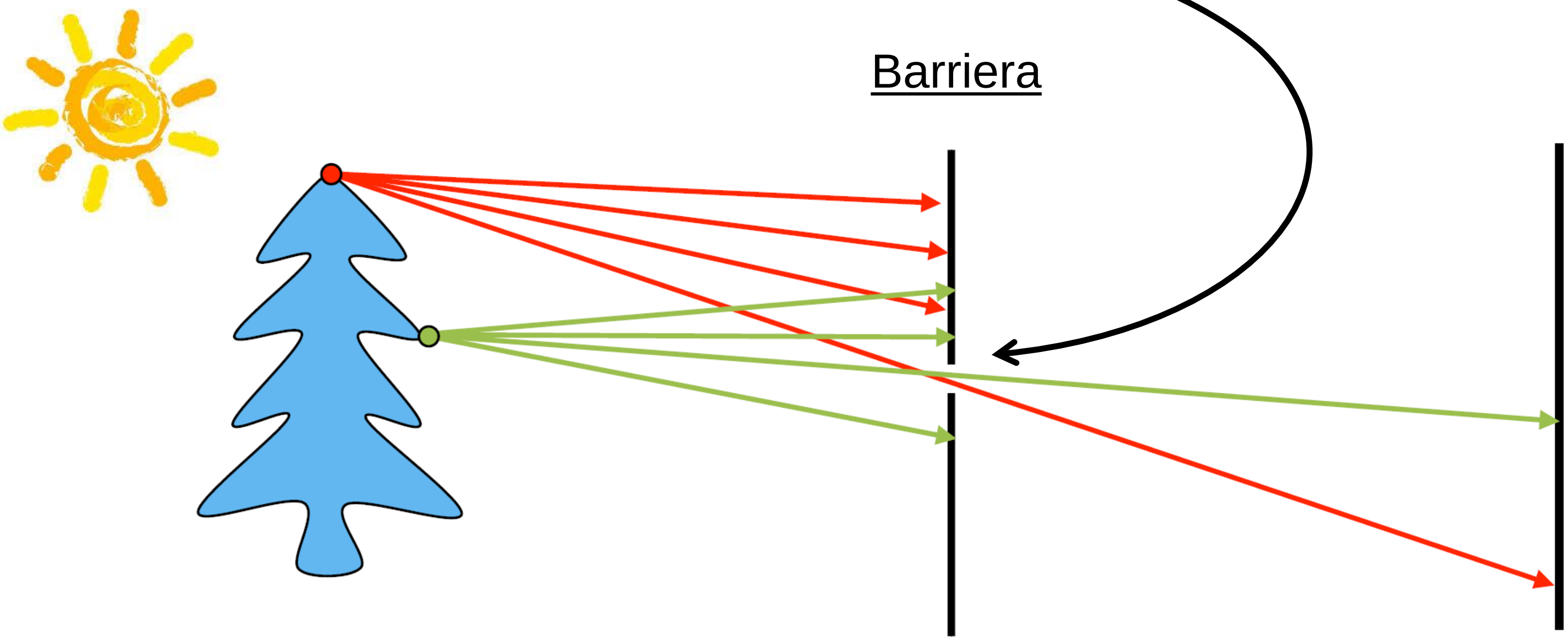
Generazione di un'immagine nelle fotocamere

- **Pellicola fotorecettiva** di fronte ad un **oggetto illuminato**



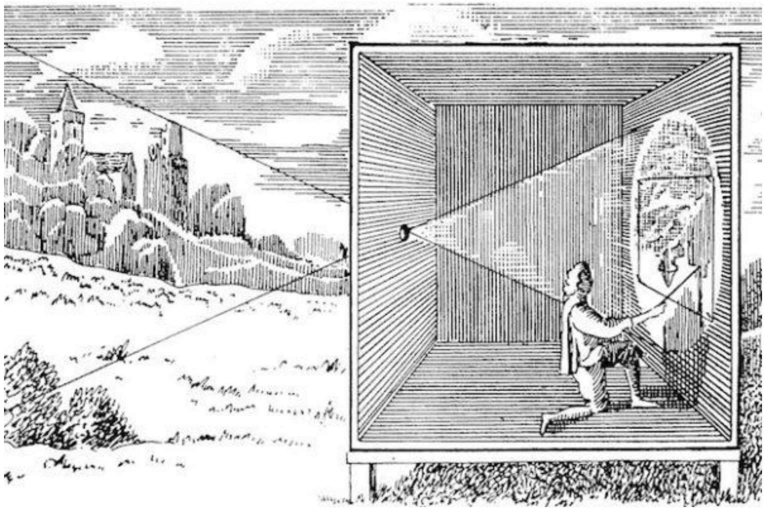
Generazione di un'immagine nelle fotocamere

- Pellicola fotorecettiva di fronte ad un oggetto illuminato
- **Barriera** per bloccare la maggior parte dei raggi
  - Riduzione della sfocatura
  - Il foro di passaggio è detto **apertura**

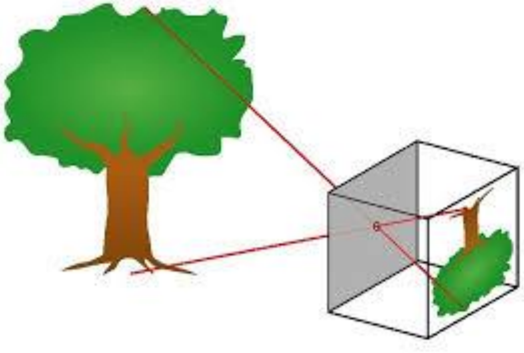


La *camera obscura*

- Principio di base già noto ai tempi di Mozi (470-390 A.C.) e Aristotele (384-322 A.C.)
- Supporto per il disegno, descritto da Leonardo da Vinci (1452-1519)

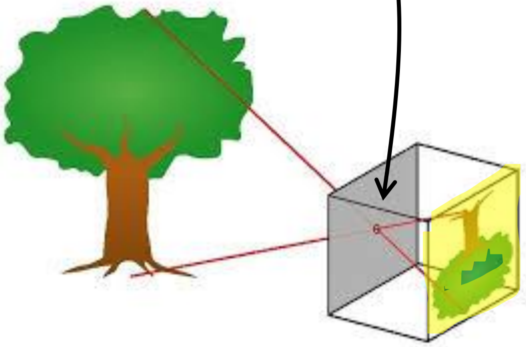


- L’immagine appare invertita
- La profondità della camera (scatola) oscura coincide con la *lunghezza focale*



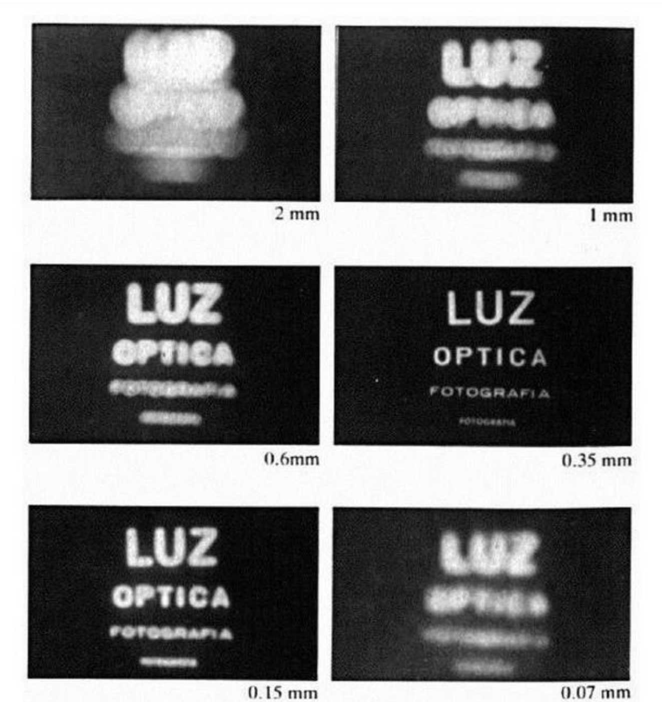
Il modello della fotocamera stenopeica

- EN: *Pinhole camera model*
- Un modello matematico per descrivere come si forma un'immagine
- Punti chiave:
  - Cattura un **fascio di raggi** – tutti i raggi attraversano lo stesso singolo punto ideale (N.B. nessuna lente!)
  - Il punto è chiamato **Centro di Proiezione** o **Centro Ottico**
  - L'immagine si forma sul **Piano dell'Immagine**



Il modello della fotocamera stenopeica

Come ridurre la sfocatura?

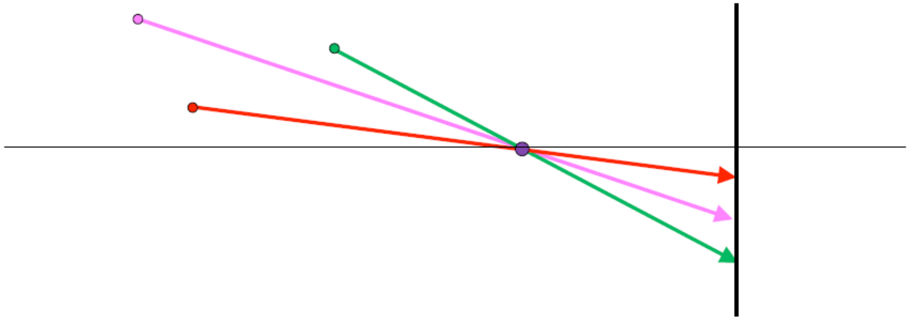


Riducendo l'apertura il più possibile?

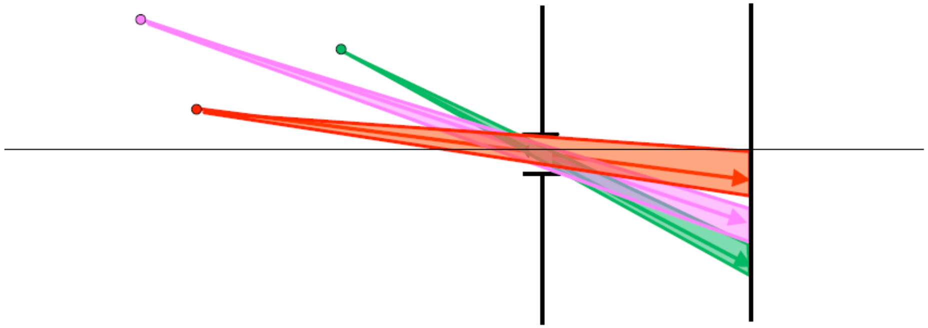
- Meno luce riesce a filtrare → bisogna aumentare il tempo di esposizione
- Effetti di diffrazione della luce!

Perché usare una lente

- Modello ideale:
  ogni punto sulla pellicola è illuminato da un singolo raggio di luce
  - Immagine fioca
  - Effetti di diffrazione

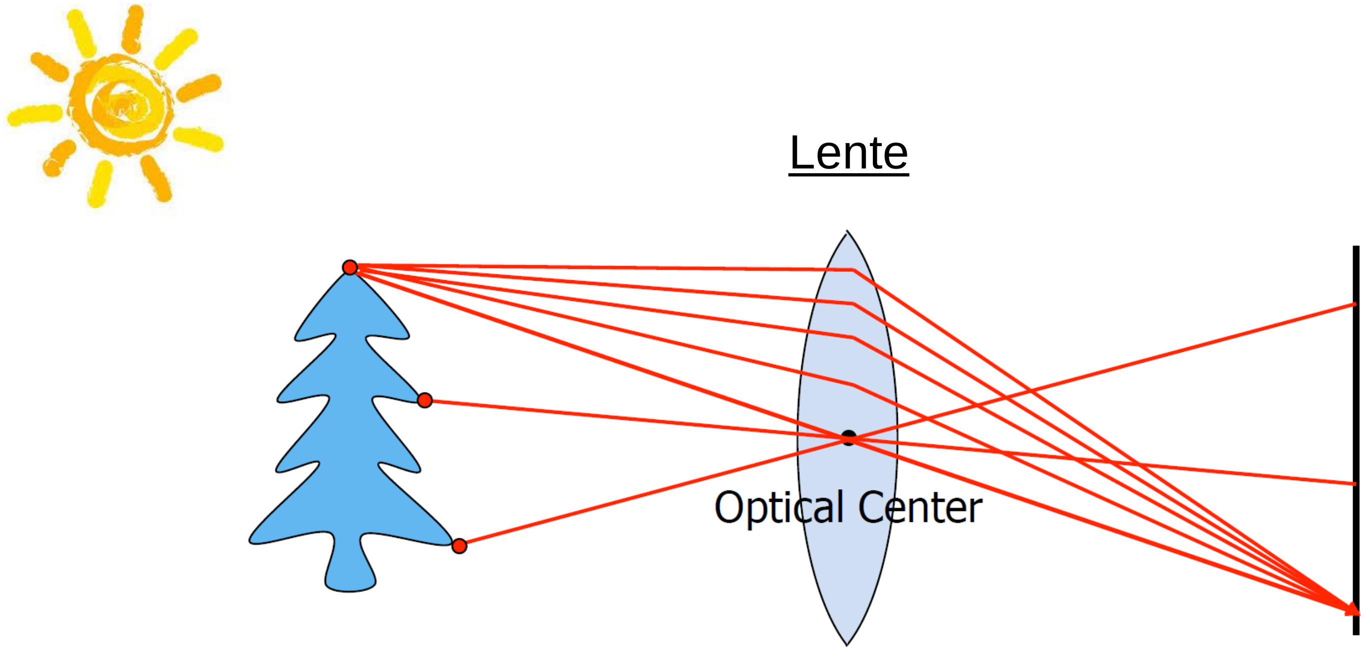


- Modello realistico:
  il foro non è infinitesimale e rende l'immagine sfocata



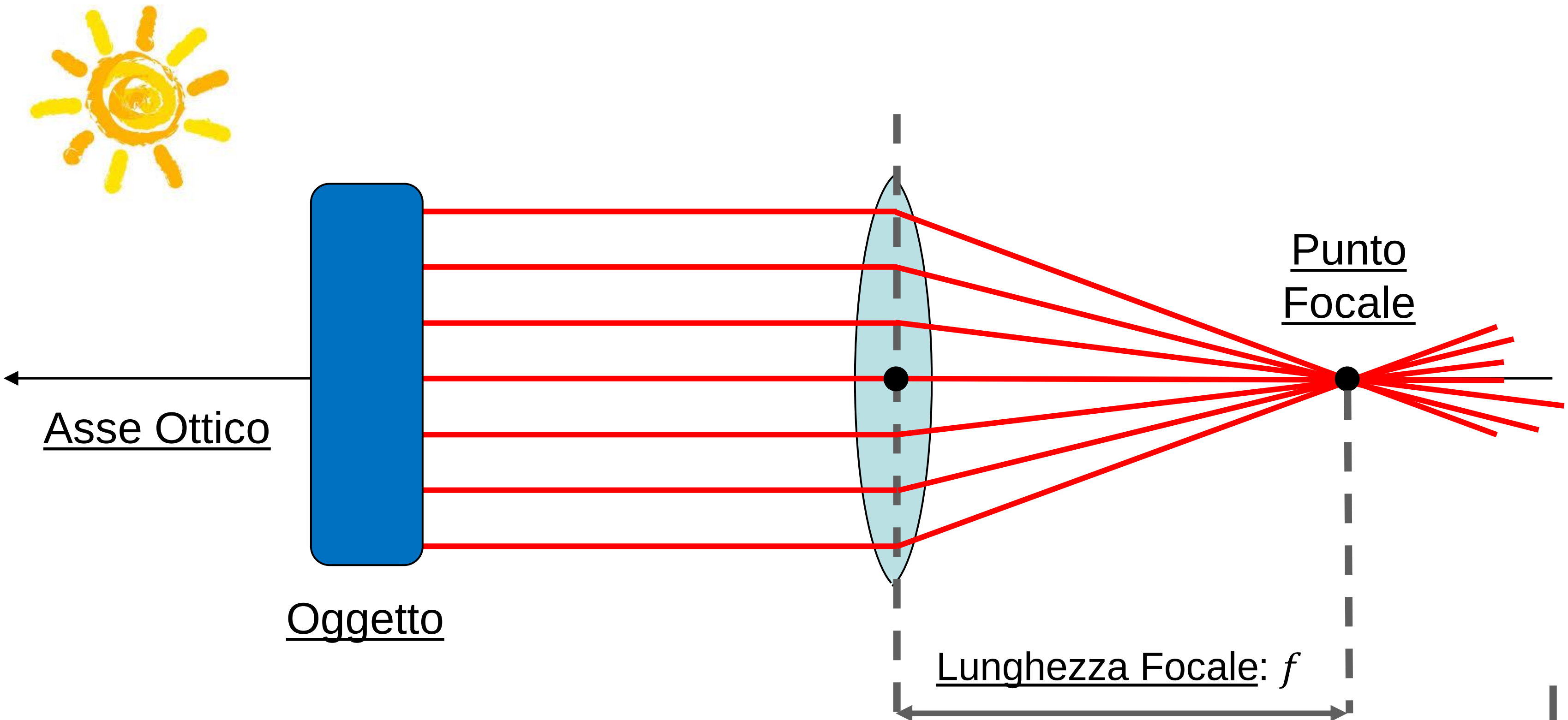
Perché usare una lente

- La lente concentra la luce sulla pellicola
- I raggi che attraversano il centro ottico non vengono deviati

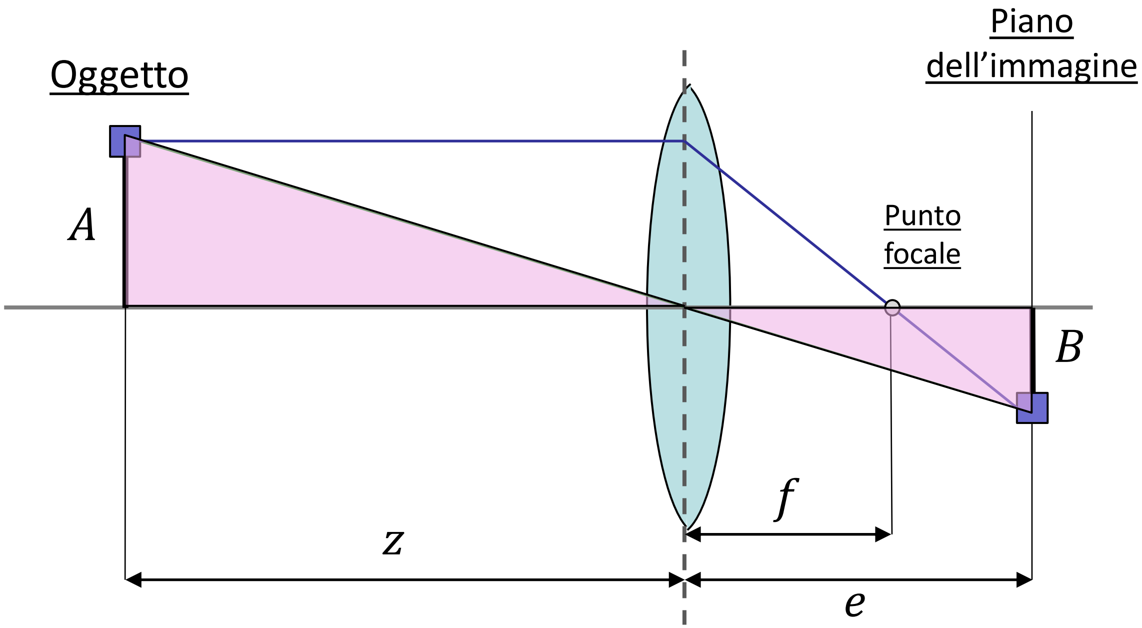


Perché usare una lente

- Tutti i raggi paralleli all'**asse ottico** convergono sul **punto focale**

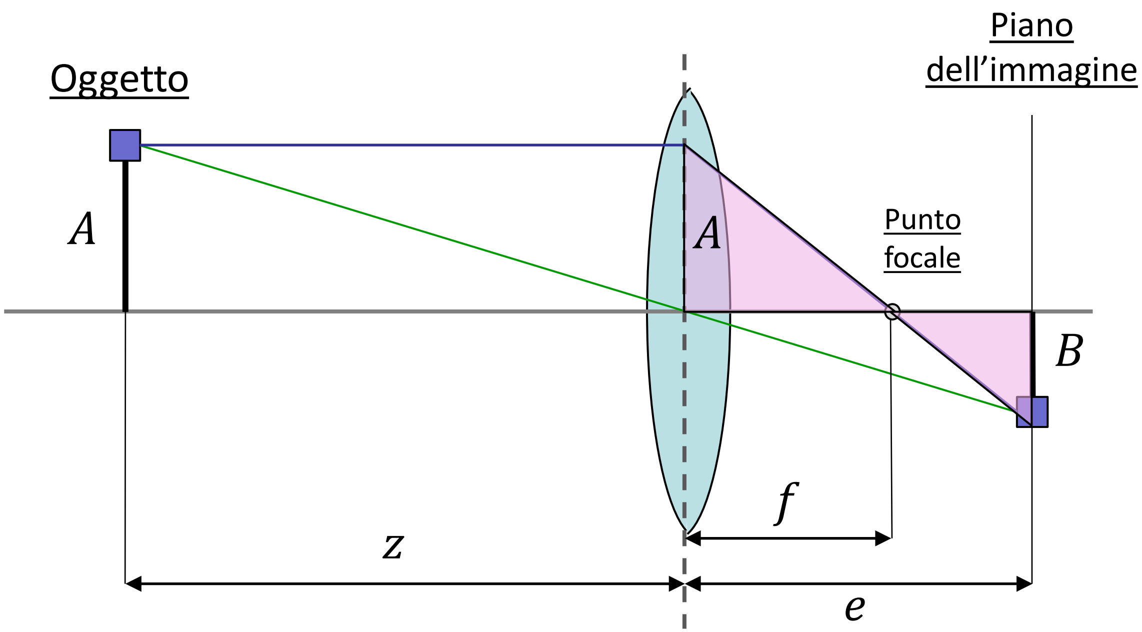


Come si genera un'immagine a fuoco



Triangoli simili: $$\frac{B}{A} = \frac{e}{z}$$

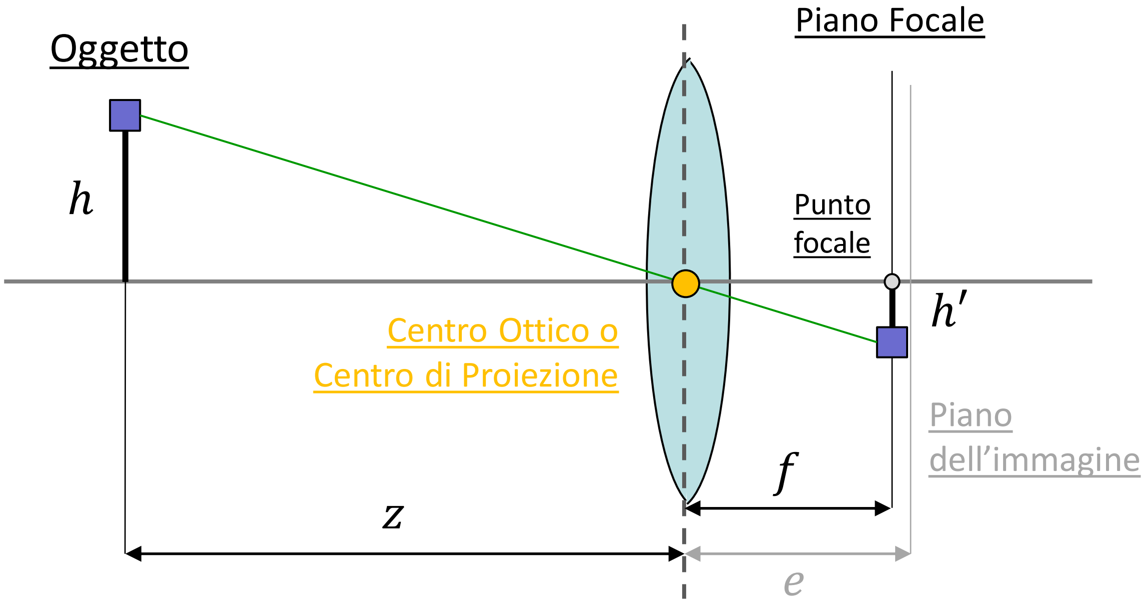
Come si genera un'immagine a fuoco



Equazione della «lente sottile»

Triangoli simili:
$$\left.\begin{array}{l} \dfrac{B}{A} = \dfrac{e}{z} \\ \dfrac{B}{A} = \dfrac{e-f}{z} = \dfrac{e}{f} - 1 \end{array}\right\} \dfrac{e}{f} - 1 = \dfrac{e}{z} \Rightarrow \boxed{\dfrac{1}{f} = \dfrac{1}{z} + \dfrac{1}{e}}$$

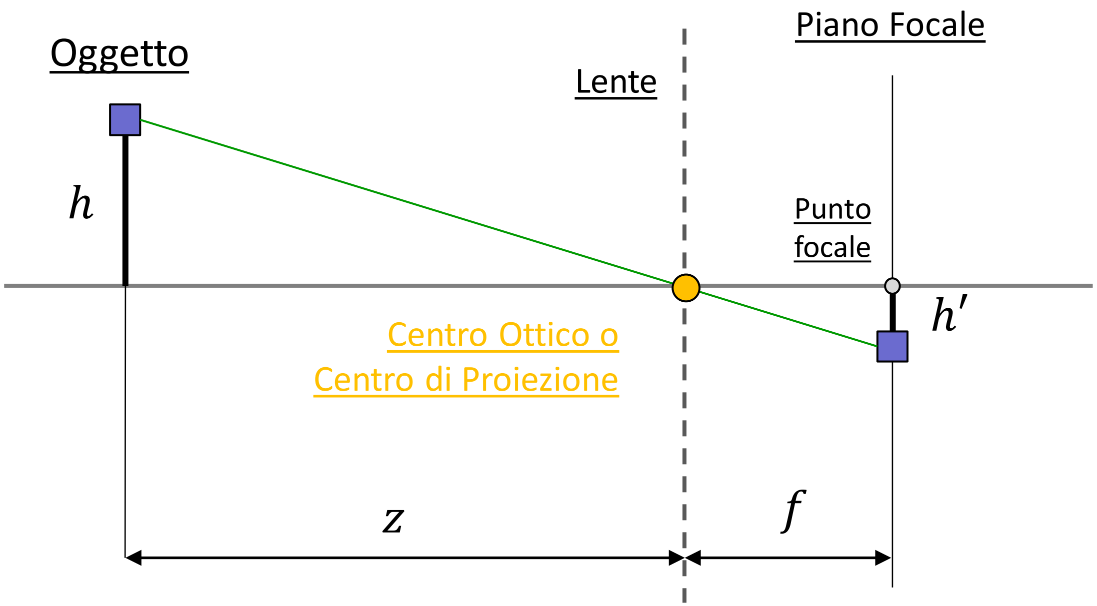
Approssimazione stenopeica



- Cosa succede per $z \gg f$?
- Aggiustiamo la posizione del piano dell'immagine in modo che gli oggetti a distanza infinita siano a fuoco

$$\frac{1}{f} = \underbrace{\boxed{\frac{1}{z}}}_{\cong 0} + \frac{1}{e} \Rightarrow \frac{1}{f} \approx \frac{1}{e} \Rightarrow f \approx e$$

Approssimazione stenopeica

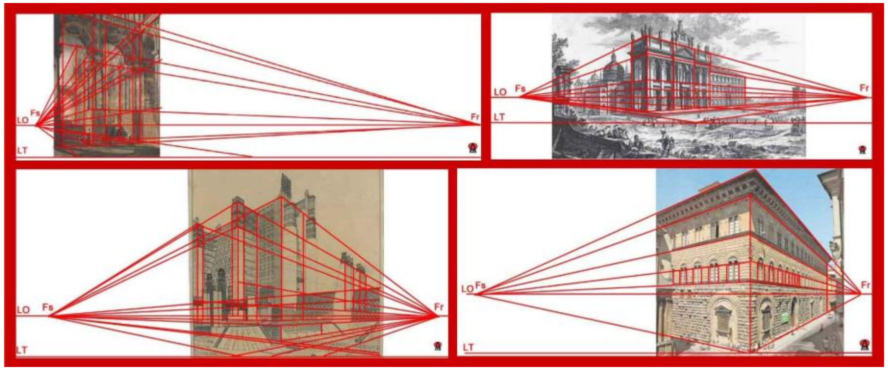


- La correlazione tra la dimensione reale di un oggetto e la sua **profondità** (i.e. la distanza dalla fotocamera) viene detta **prospettiva**

$$\frac{h'}{h} = \frac{f}{z} \Rightarrow h' = \frac{f}{z} h$$

Effetti prospettici

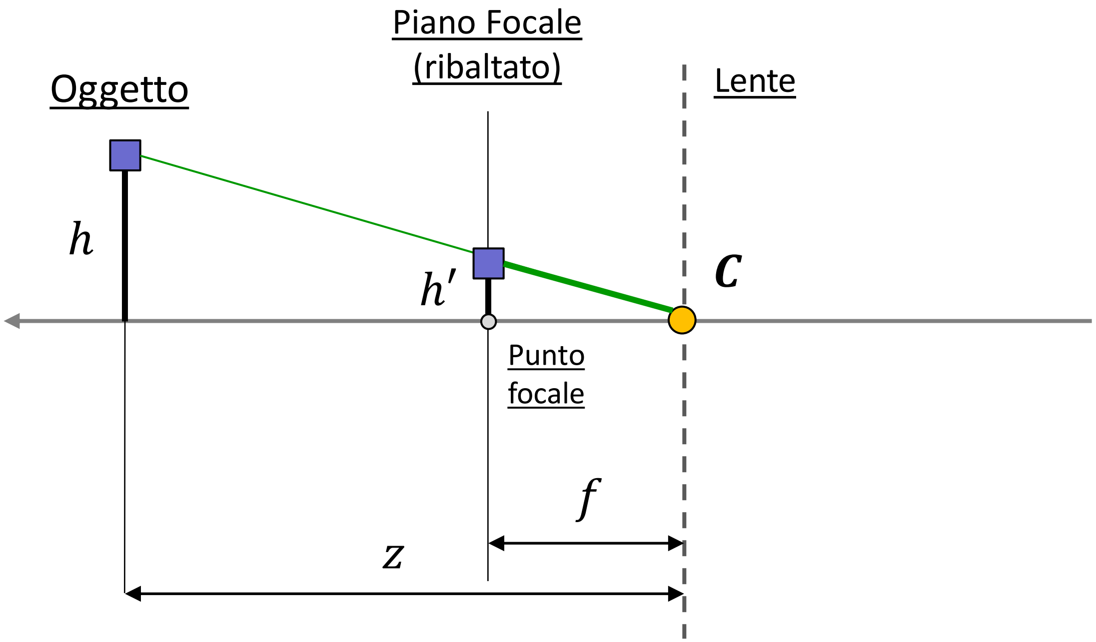
- Grazie alla prospettiva siamo in grado di interpretare la distanza degli oggetti
  - Percepire una scena 3D dalla sua rappresentazione 2D (immagine)
- Oggetti lontani appaiono più piccoli
- Linee parallele nel mondo si intersecano nel **punto di fuga** dell’immagine



Linee di Fuga

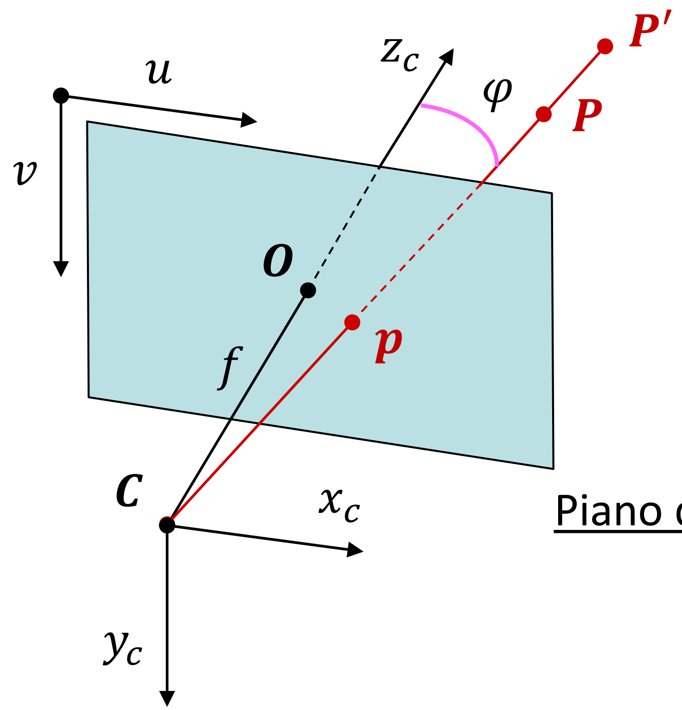
Proiezione prospettica

- Per convenienza, il piano dell'immagine è solitamente rappresentato di fronte al centro ottico (i.e. $\boldsymbol{C}$) in modo da preservare lo stesso orientamento della scena inquadrata (non capovolta, quindi)



Proiezione prospettica

- La fotocamera (2D) **non misura le distanza**, ma gli angoli!
  → sensore di orientamento



z_c = asse ottico

$\boldsymbol{O}$ = punto principale

$\boldsymbol{C}$ = centro ottico
= centro della lente

Piano dell'immagine

Proiezione prospettica: dal mondo ai pixel

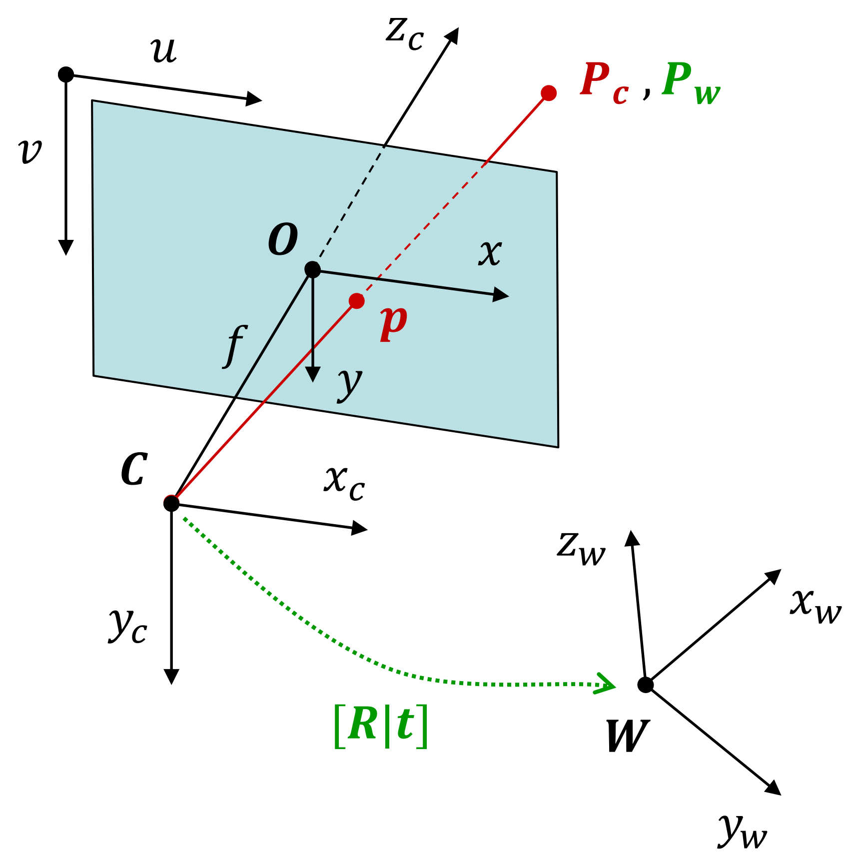
Trovare le coordinate (u, v) del punto $\boldsymbol{P_w}$ nel sistema di riferimento globale $(\boldsymbol{W})$

0. Convertire il punto $\boldsymbol{P_c}$ espresso nel sistema globale in $\boldsymbol{P_w}$, nel sistema di riferimento della fotocamera

Trovare le coordinate (u, v) del punto $\boldsymbol{P_c}$ nel sistema di riferimento della telecamera $(\boldsymbol{C})$

1. Convertire il punto $\boldsymbol{P_c}$ in coordinate (x, y) nel piano dell'immagine

2. Convertire il punto $\boldsymbol{p}$ in coordinate (discrete) (u, v) nello spazio dei pixel



Proiezione prospettica: dalla fotocamera al piano dell'immagine

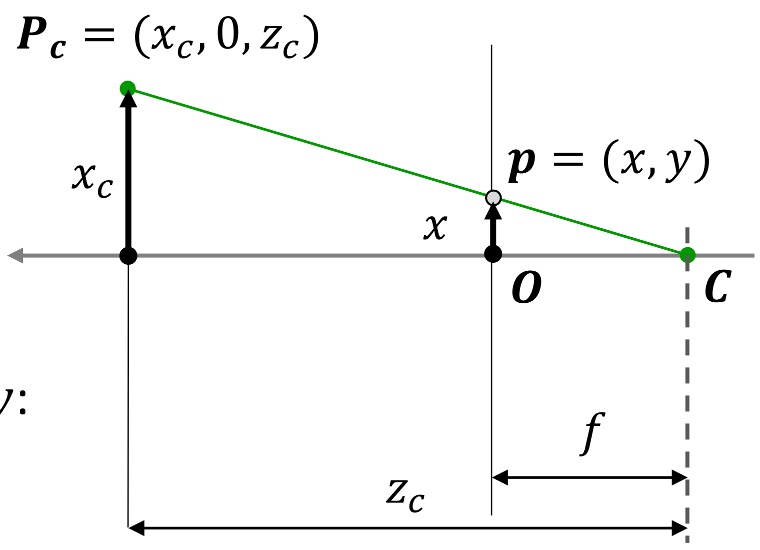
- Il punto (tridimensionale) espresso in «coordinate fotocamera» $\boldsymbol{P_c} = (x_c, 0, z_c)$ è proiettato sul punto (bidimensionale) $\boldsymbol{p} = (x, y)$ sul piano delle immagini

- Sfruttando triangoli simili:

$$\frac{x}{f} = \frac{x_c}{z_c} \Rightarrow x = \frac{f x_c}{z_c}$$

- Allo stesso modo, per le coordinate y:

$$\frac{y}{f} = \frac{y_c}{z_c} \Rightarrow y = \frac{f y_c}{z_c}$$

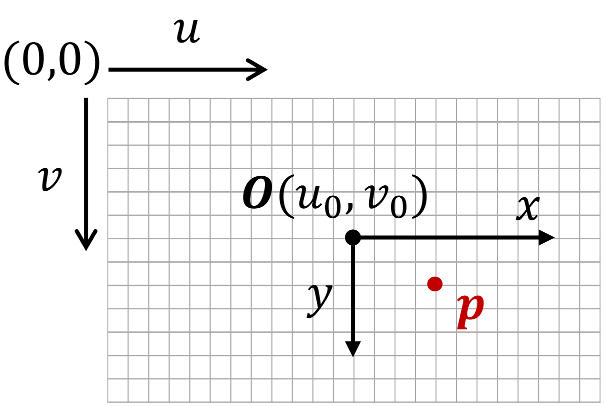


Proiezione prospettica: dalla fotocamera ai pixels

- Per convertire $\boldsymbol{p}$ da coordinate locali nel piano delle immagini (x, y) alle coordinate pixel (u, v) bisogna considerare:
  - Le coordinate in pixels del centro ottico della fotocamera $\boldsymbol{O} = (u_0, v_0)$
  - I fattori di scala k_u, k_v per passare da *m* a *pixels* in entrambe le dimensioni
- Risulta quindi:

$$u = u_0 + k_u x \;\Rightarrow u = u_0 + \frac{k_u f x_c}{z_c}$$

$$v = v_0 + k_v y \;\Rightarrow v = v_0 + \frac{k_v f y_c}{z_c}$$



- **Coordinate Omogenee** per una mappatura lineare da 3D a 2D, introducendo un elemento extra (scala) λ

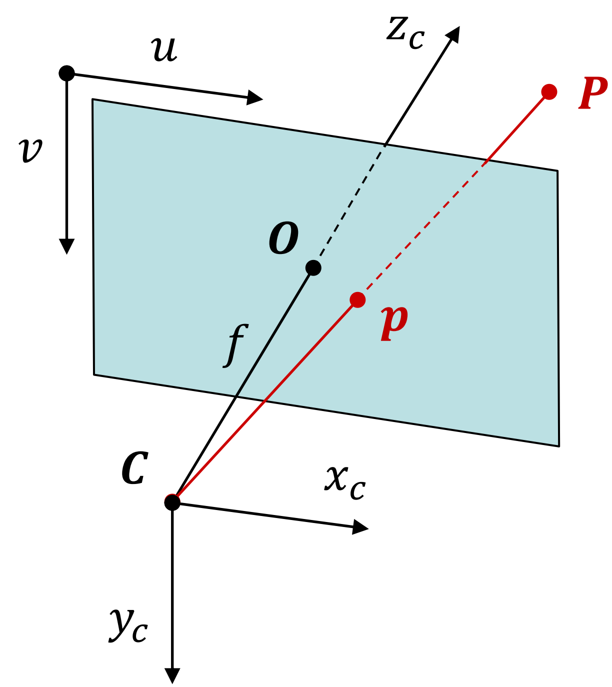
$$\boldsymbol{p} = \binom{u}{v} \quad \Longrightarrow \quad \tilde{\boldsymbol{p}} = \begin{bmatrix} \lambda u \\ \lambda v \\ \lambda \end{bmatrix} = \lambda \begin{bmatrix} u \\ v \\ 1 \end{bmatrix}$$

N.B. solitamente λ = 1

Proiezione prospettica: dalla fotocamera ai pixels

- In forma matriciale utilizzando le coordinate omogenee:

$$\begin{bmatrix}\lambda u\\ \lambda v\\ \lambda\end{bmatrix} = \begin{bmatrix}k_u f & 0 & u_0\\ 0 & k_v f & v_0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}x_c\\ y_c\\ z_c\end{bmatrix}$$



- O alternativamente:

$$\begin{bmatrix}\lambda u\\ \lambda v\\ \lambda\end{bmatrix} = \begin{bmatrix}\alpha_u & 0 & u_0\\ 0 & \alpha_v & v_0\\ 0 & 0 & 1\end{bmatrix}\begin{bmatrix}x_c\\ y_c\\ z_c\end{bmatrix} = \boldsymbol{K}\begin{bmatrix}x_c\\ y_c\\ z_c\end{bmatrix}$$

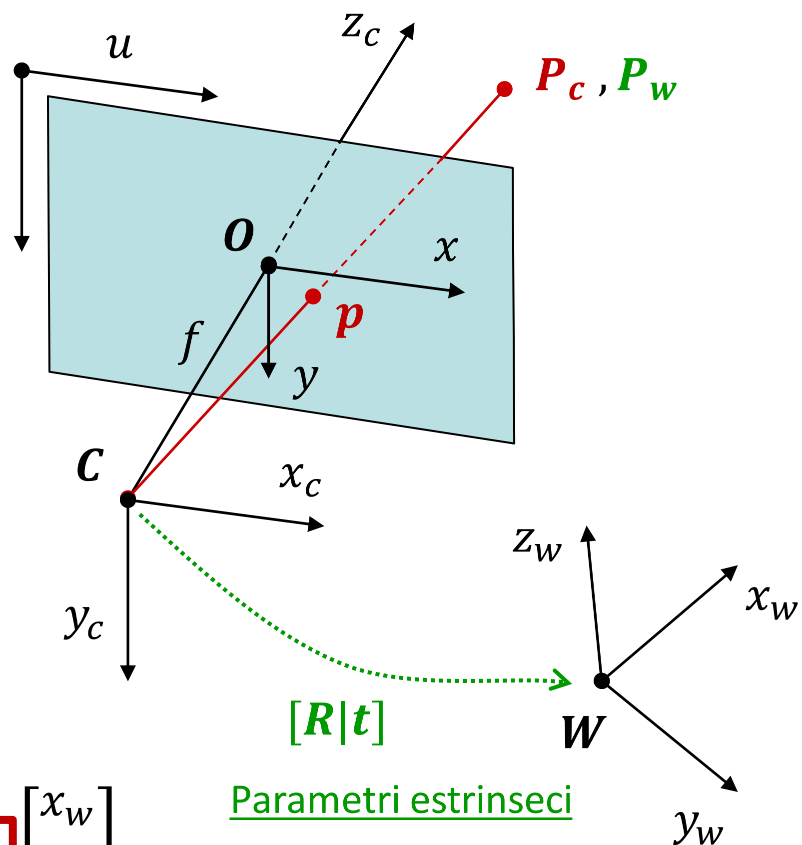
- α_u: lunghezza focale nella direzione u
- α_v: lunghezza focale nella direzione v
- $\boldsymbol{K}$: «Matrice di calibrazione» o «Matrice dei **parametri intrinseci**»

Proiezione prospettica: dal mondo alla fotocamera

$$\begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} + \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix}$$

$$\begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} = \left[\begin{array}{ccc|c} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{array}\right] \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}$$

$$= \left[\begin{array}{ccc|c} & \boldsymbol{R} & & \boldsymbol{t} \\ 0 & 0 & 0 & 1 \end{array}\right] \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}$$



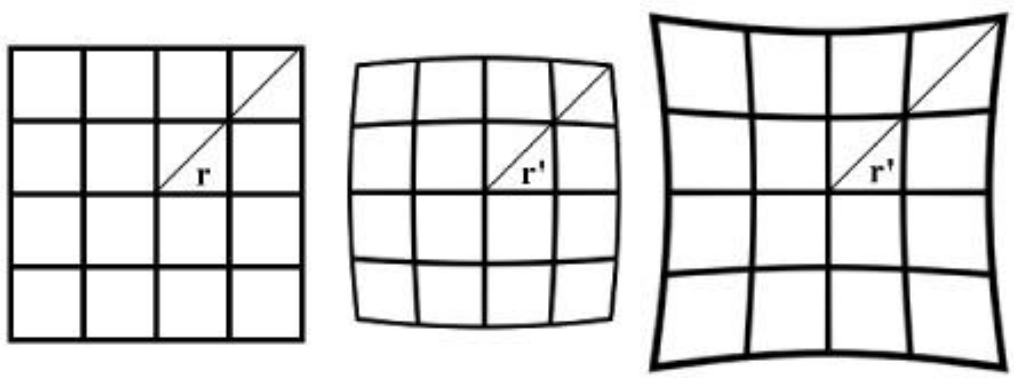
Matrice di Proiezione

Dalla slide precedente si ottiene:

$$\lambda \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \boxed{\boldsymbol{K}[\boldsymbol{R}|\boldsymbol{t}]} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}$$

Proiezione prospettica: distorsione radiale

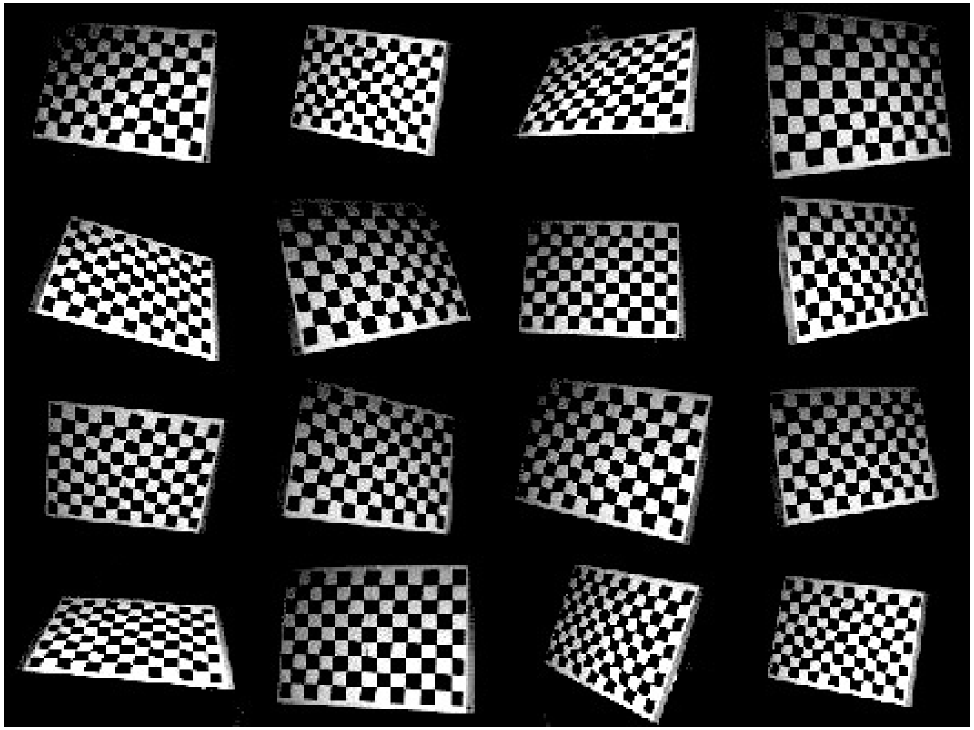
- La distorsione è una funzione non lineare della distanza dal centro dell'immagine
- Dalle coordinate ideali (u, v) a quelle distorte (u_d, v_d):
  - Modello quadratico di distorsione
    $$\begin{bmatrix} u_d \\ v_d \end{bmatrix} = (1 + k_1 r^2) \begin{bmatrix} u - u_0 \\ v - v_0 \end{bmatrix} + \begin{bmatrix} u_0 \\ v_0 \end{bmatrix}$$
    dove: $r^2 = (u - u_0)^2 + (v - v_0)^2$
  - Buona approssimazione per gran parte delle lenti



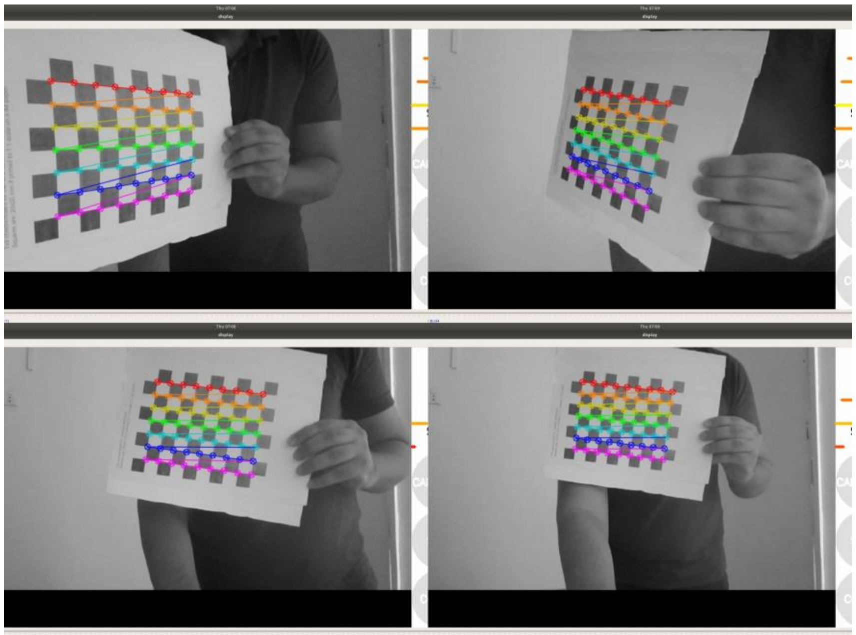
No distorsione | Distorsione «a barile» | Distorsione «a cuscino»

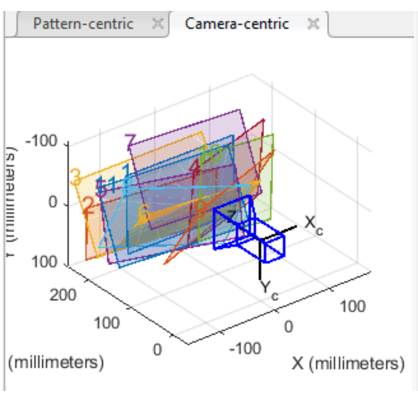
Calibrazione di una fotocamera

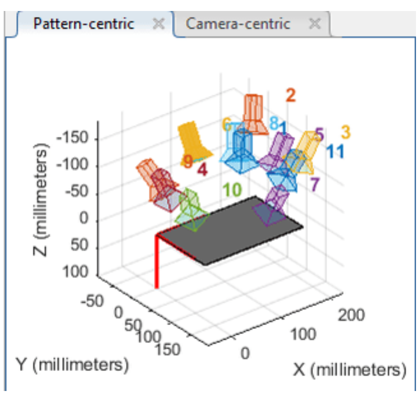
- Necessaria per trovare i *parametri intrinseci* di una telecamera
- Necessaria per passare dal mondo 3D a quello 2D e viceversa



Calibrazione di una fotocamera







Come funziona:

- Modello della fotocamera stenopeica con distorsione per interpretare la proiezione mondo ←→ immagine
- Pattern di dimensione e forma note, facilmente individuabile nell'immagine
- Conoscendo la corrispondenza $\boldsymbol{p} \leftrightarrow \boldsymbol{P}$, è possibile calcolare i parametri ignoti $\boldsymbol{K}, \boldsymbol{R}, \boldsymbol{t}$ applicando l'equazione prospettica
- **EXTRA:** è possibile anche stimare le distanze e pose fisiche del pattern rispetto alla telecamera («parametri estrinseci» della calibrazione)

Calibrazione di una fotocamera

Metodo:

1. *Trasformazione Lineare Diretta (TLD)**
2. Si considera il sistema globale fisso sul pattern scelto (di solito l'angolo in alto a sinistra) → posizioni 3D degli elementi del pattern note
3. Sapendo che: $\lambda \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \boldsymbol{K}[\boldsymbol{R}|\boldsymbol{t}] \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}$
4. **11 valori** da stimare (m_{34} può essere considerato = 1, poiché il fattore di scala non ci interessa), **2 equazioni** per ogni corrispondenza tra punto 3D e sua proiezione 2D → servono **almeno 6 corrispondenze** in un'immagine → metodo dei minimi quadrati (lineare)
5. *Decomposizione QR* per decomporre la sottomatrice 3x3 ($m_{11}: m_{33}$) nel prodotto di una matrice triangolare superiore ($\boldsymbol{K}$) e una di rotazione ($\boldsymbol{R}$) (ortogonale)
6. Traslazione: $\boldsymbol{t} = \boldsymbol{K}^{-1}[m_{14} \quad m_{24} \quad m_{34}]^T$

* Zhang, Z. "A Flexible New Technique for Camera Calibration." *IEEE Transactions on Pattern Analysis and Machine Intelligence*. Vol. 22, No. 11, 2000, pp. 1330–1334.

ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Prof. Marco Carricato

Prof. Edoardo Idà

Ing. Simone Comari

DIN

marco.carricato@unibo.it

edoardo.ida2@unibo.it

simone.comari@unibo.it

https://www.unibo.it/sitoweb/marco.carricato/

https://www.unibo.it/sitoweb/edoardo.ida2/

https://www.unibo.it/sitoweb/simone.comari/