

رسالة محمد



مبانی بینایی کامپیوتر

مدرس: محمدرضا محمدی

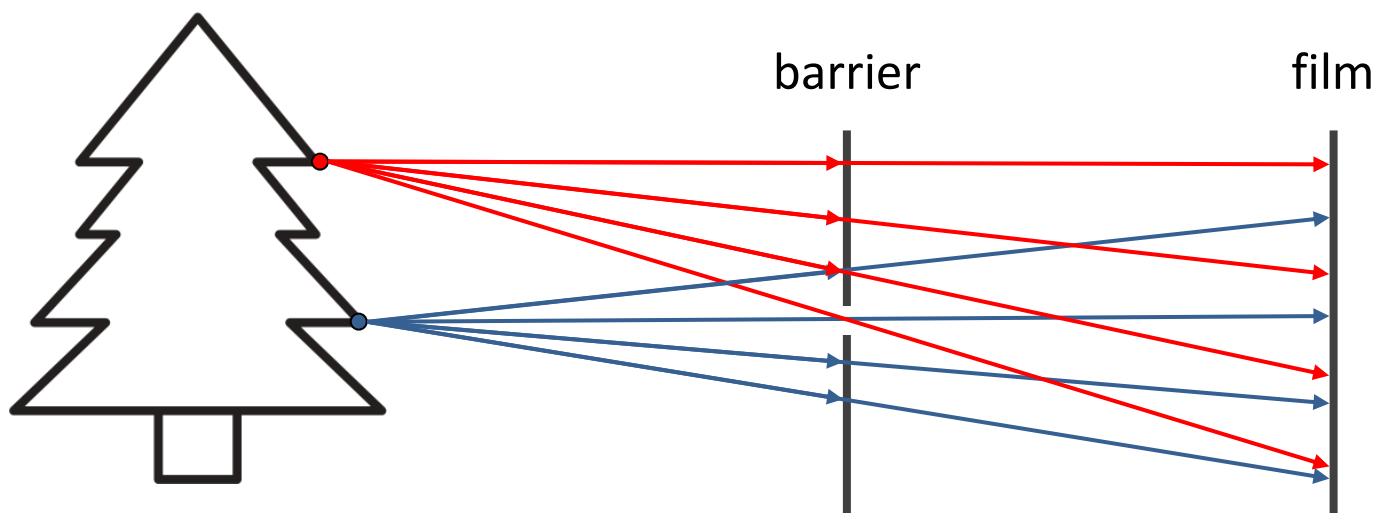
۱۴۰۱

تشکیل تصویر

Image Formation

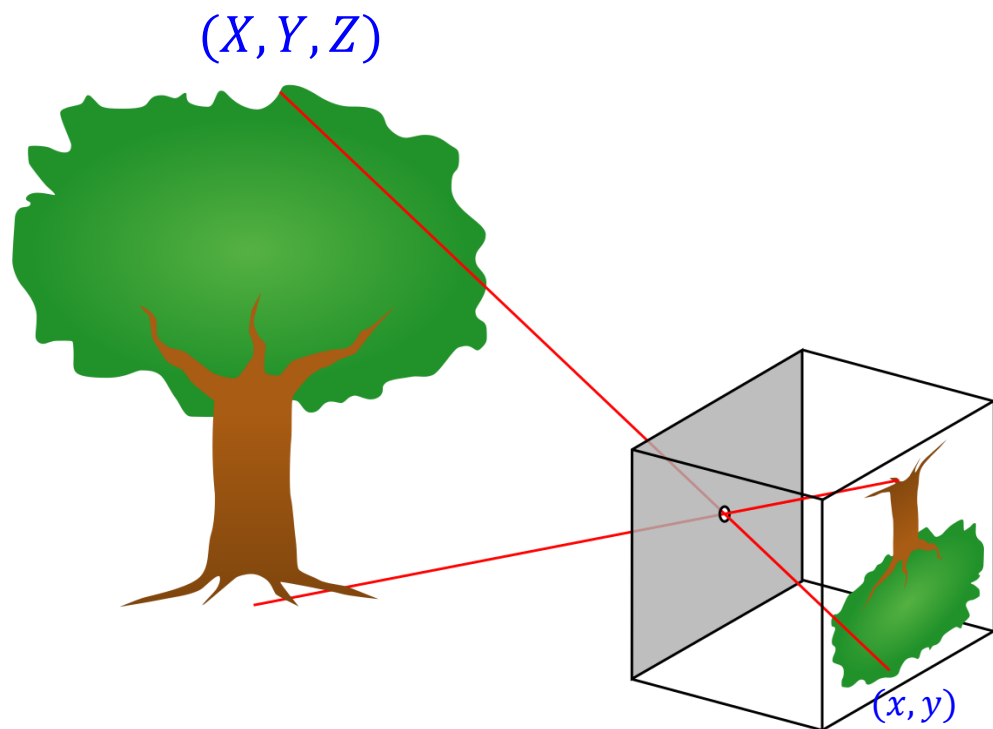
طراحی دوربین

- فرض کنید یک فیلم را مقابل یک شیء قرار دهیم
- باید مانعی (دریچه‌ای) در مقابل حسگرها قرار دهیم تا هر کدام نسبت به بخشی از فضا حساس باشند



مدل دوربین Pinhole

- ساده‌ترین دستگاهی است که یک تصویر از صحنه سه بعدی روی یک صفحه دو بعدی تشکیل می‌دهد



perspective projection:

$$x = f \frac{X}{Z} \quad y = f \frac{Y}{Z}$$

f : فاصله کانونی

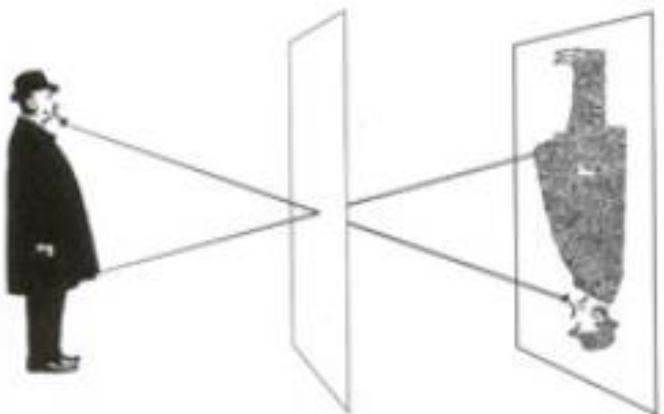
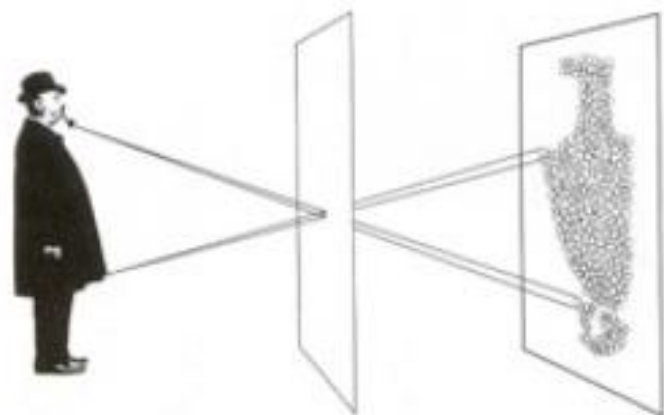
اثر اندازه دریچه

- دریچه بزرگ

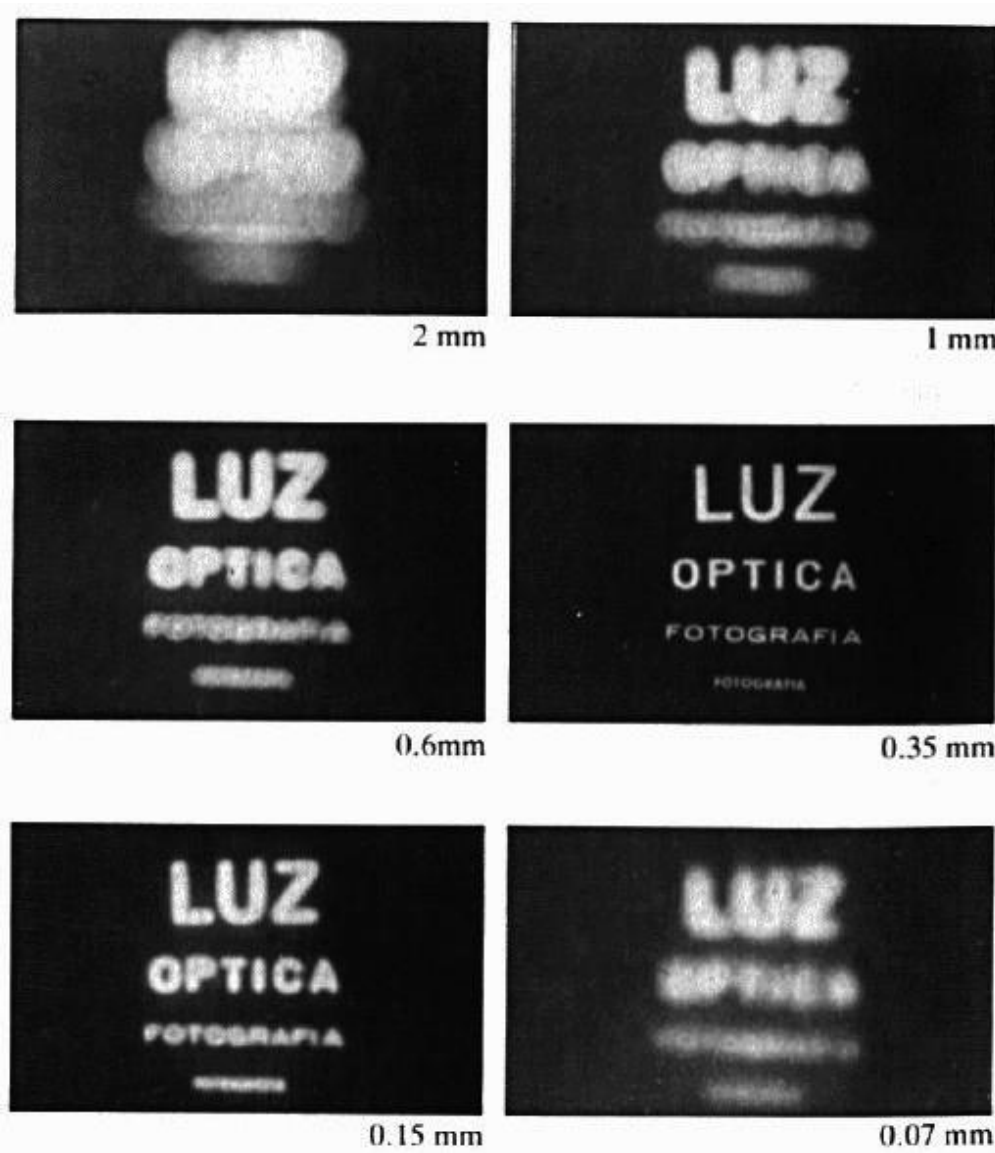
- نور منعکس شده در بخش بیشتری از تصویر اثر می گذارد
- تصویر تار خواهد بود

- دریچه کوچک

- تار شدن را کاهش می دهد اما مقدار نور وارد شده به دوربین را کم می کند
- همچنین باعث پراکندگی نور می شود

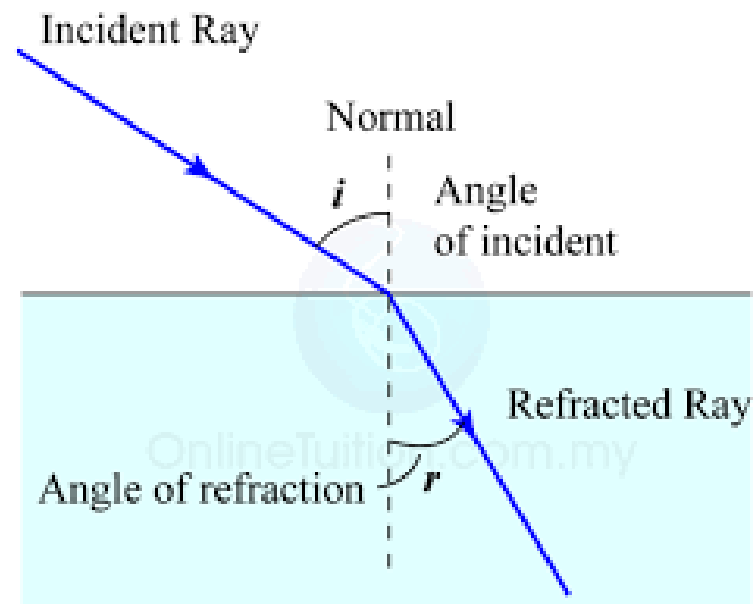
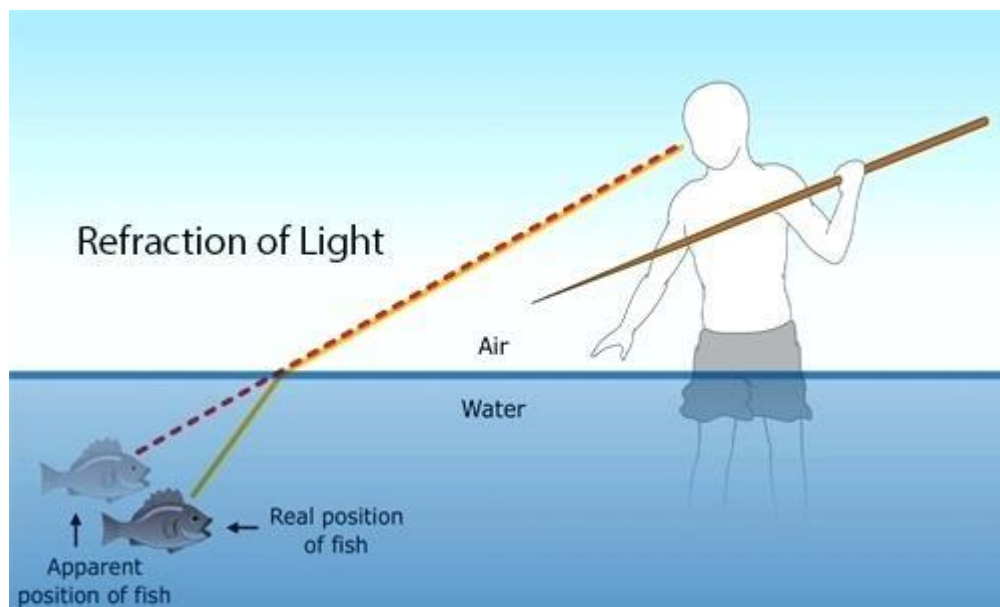


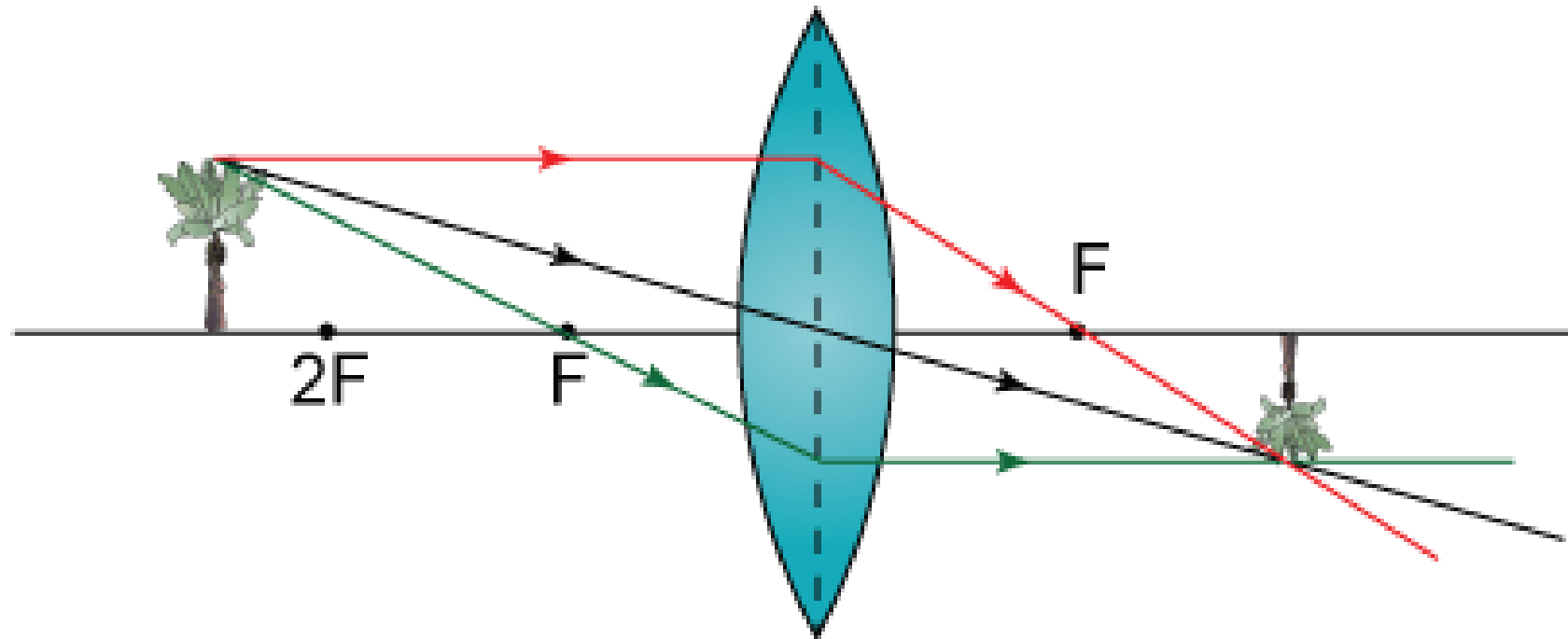
اثر اندازه دریچه

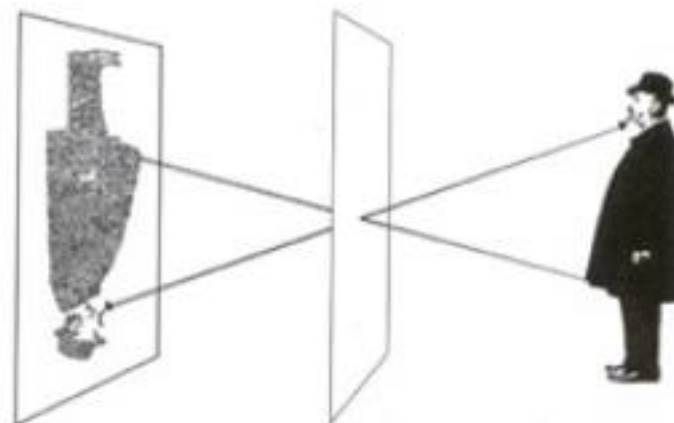
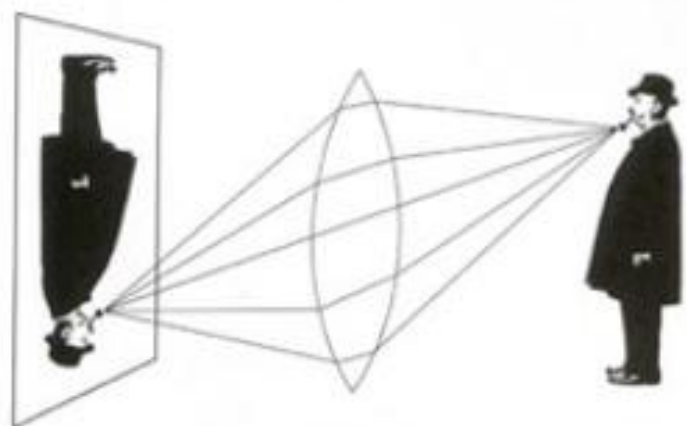


شکست نور

- خم شدن یا شکست موج هنگامی که وارد ماده‌ای با سرعت متفاوت می‌شود

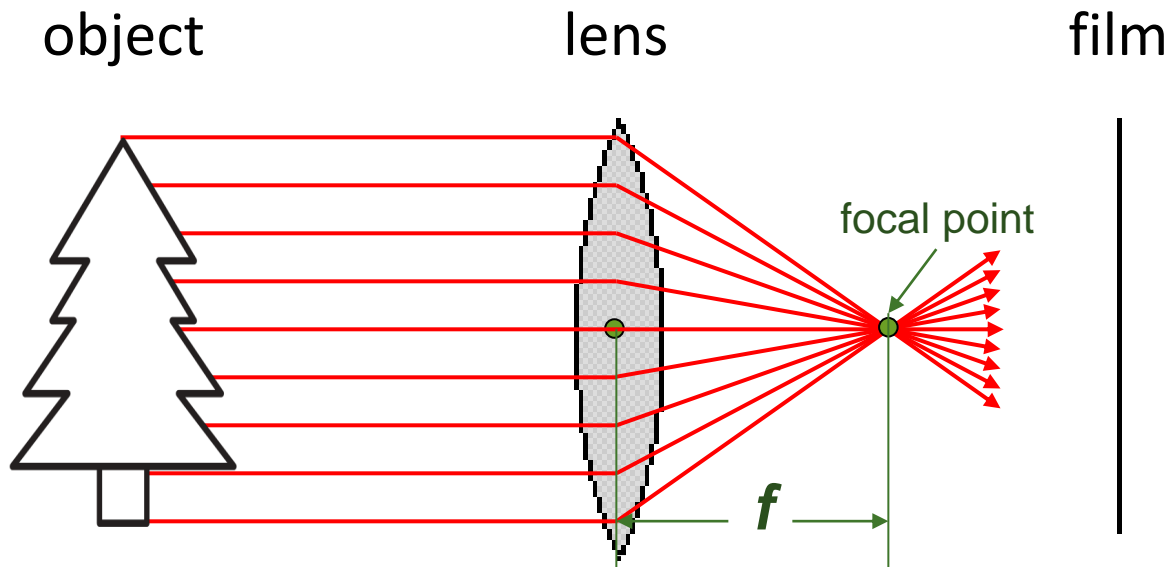






خواص لنز نازک (ایده آل)

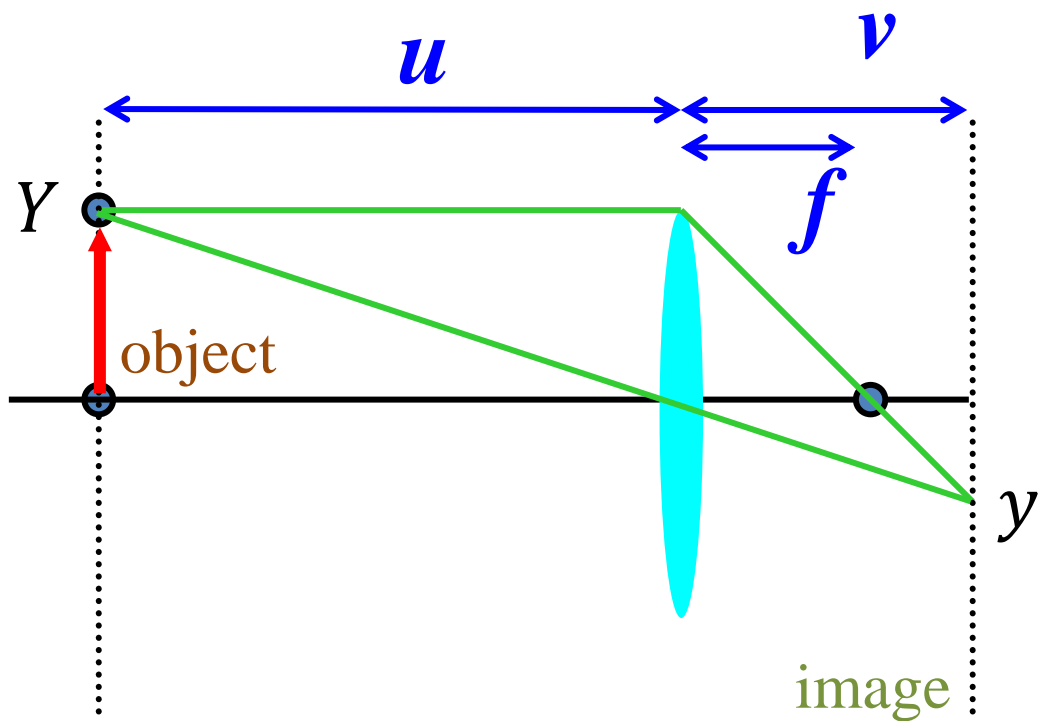
- اشعه‌های نوری که از مرکز لنز عبور می‌کنند منحرف نمی‌شوند
- میزان انحراف با دور شدن از مرکز لنز بیشتر می‌شود
- تمام خطوط موازی به یک نقطه همگرا می‌شوند



aparat.com/mostafa_karimi

معادلات لنز نازک

- فرض کنید یک شیء در فاصله u از لنز قرار دارد



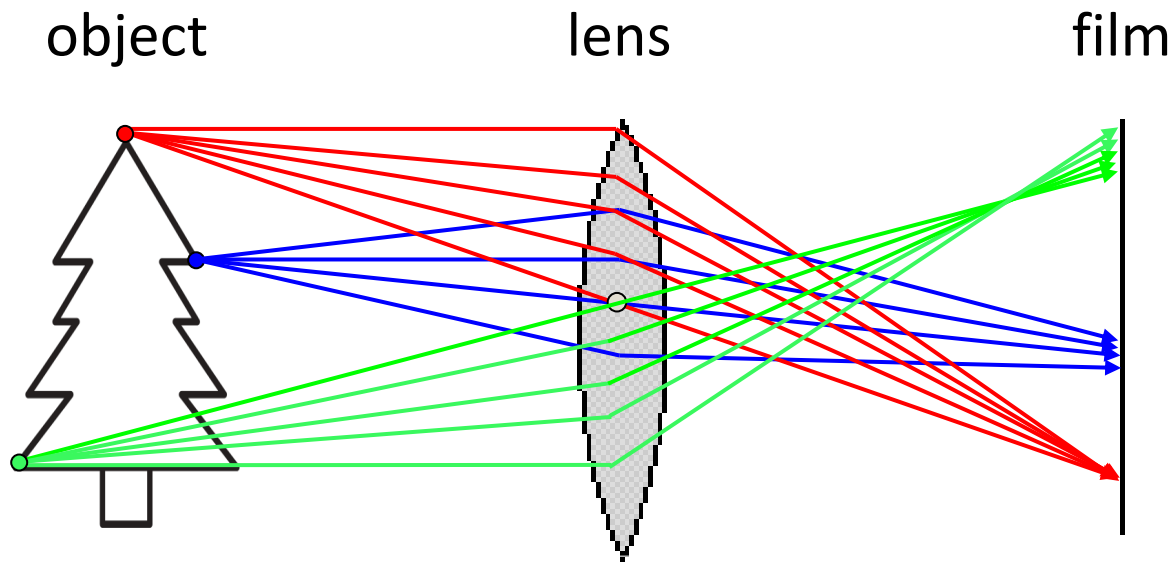
$$\frac{y}{Y} = \frac{u}{v}$$

$$\frac{y}{Y} = \frac{v - f}{f}$$

$$\frac{1}{f} = \frac{1}{v} + \frac{1}{u}$$

معادلات لنز نازک

- تنها اشعه‌های نوری نقطه‌ای که در فاصله u از لنز باشند در صفحه‌ای به فاصله v از لنز همگرا (متمرکز) می‌شوند
- نقاط با فاصله‌های دیگر دچار تاری خواهند شد



$$\frac{1}{f} = \frac{1}{v} + \frac{1}{u}$$

عمق میدان (DOF)

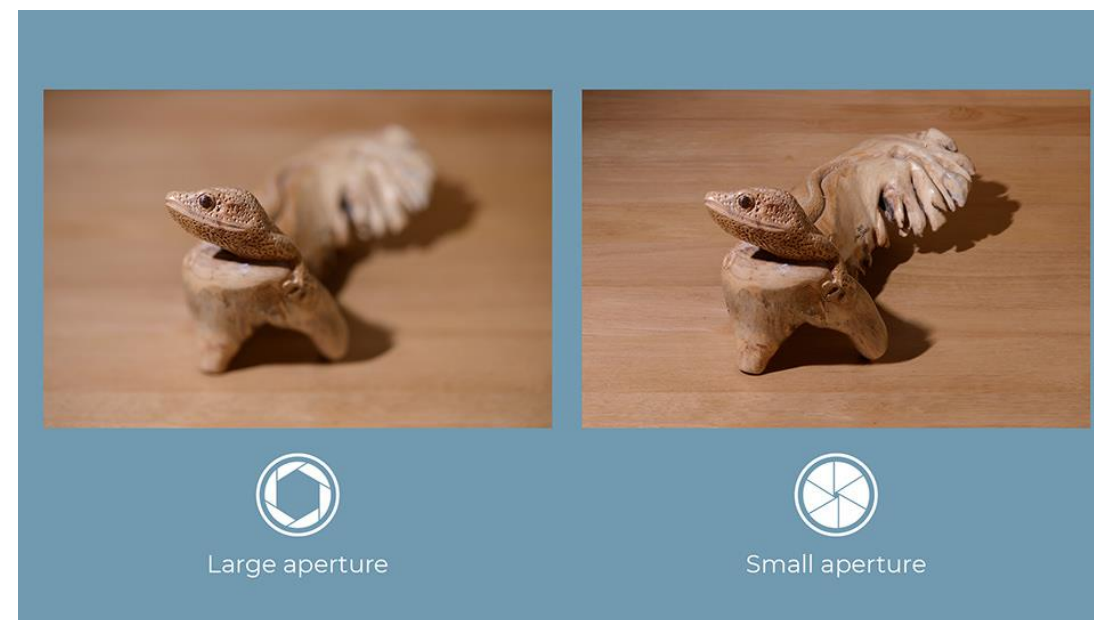
- محدوده‌ای از عمق (فاصله تا دوربین) که اشیاء تقریباً با وضوح مناسب دیده می‌شوند



عمق میدان (DOF)

- در دوربین‌ها معمولاً هم از لنز استفاده می‌شود و هم از دریچه استفاده می‌شود و می‌توان عمق میدان را کنترل کرد

	Aperture Size	Exposure	Depth of Field
f/1.4	Very large	Lets in a lot of light	Very thin
f/2.0	Large	Half as much light as f/1.4	Thin
f/2.8	Large	Half as much light as f/2	Thin
f/4.0	Moderate	Half as much light as f/2.8	Moderately thin
f/5.6	Moderate	Half as much light as f/4	Moderate
f/8.0	Moderate	Half as much light as f/5.6	Moderately large
f/11.0	Small	Half as much light as f/8	Large
f/16.0	Small	Half as much light as f/11	Large
f/22.0	Very small	Half as much light as f/16	Very large

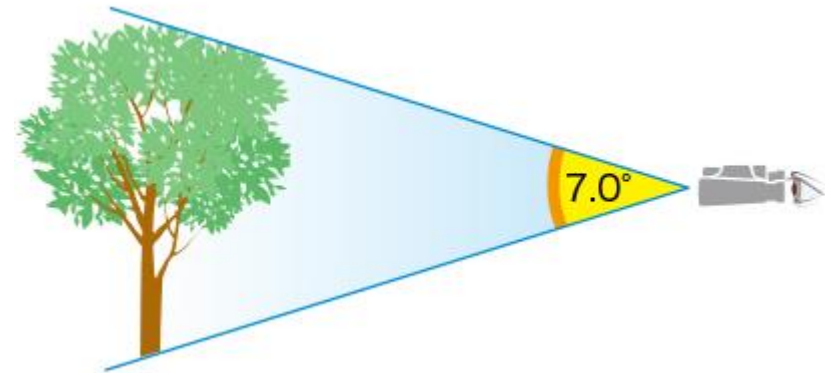
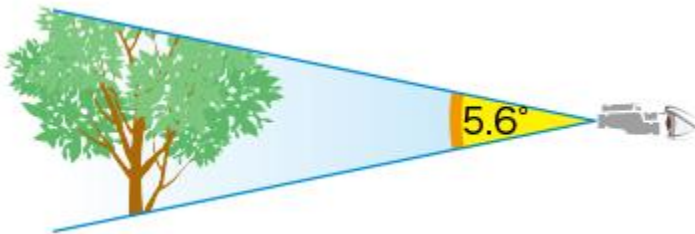


Large aperture

Small aperture

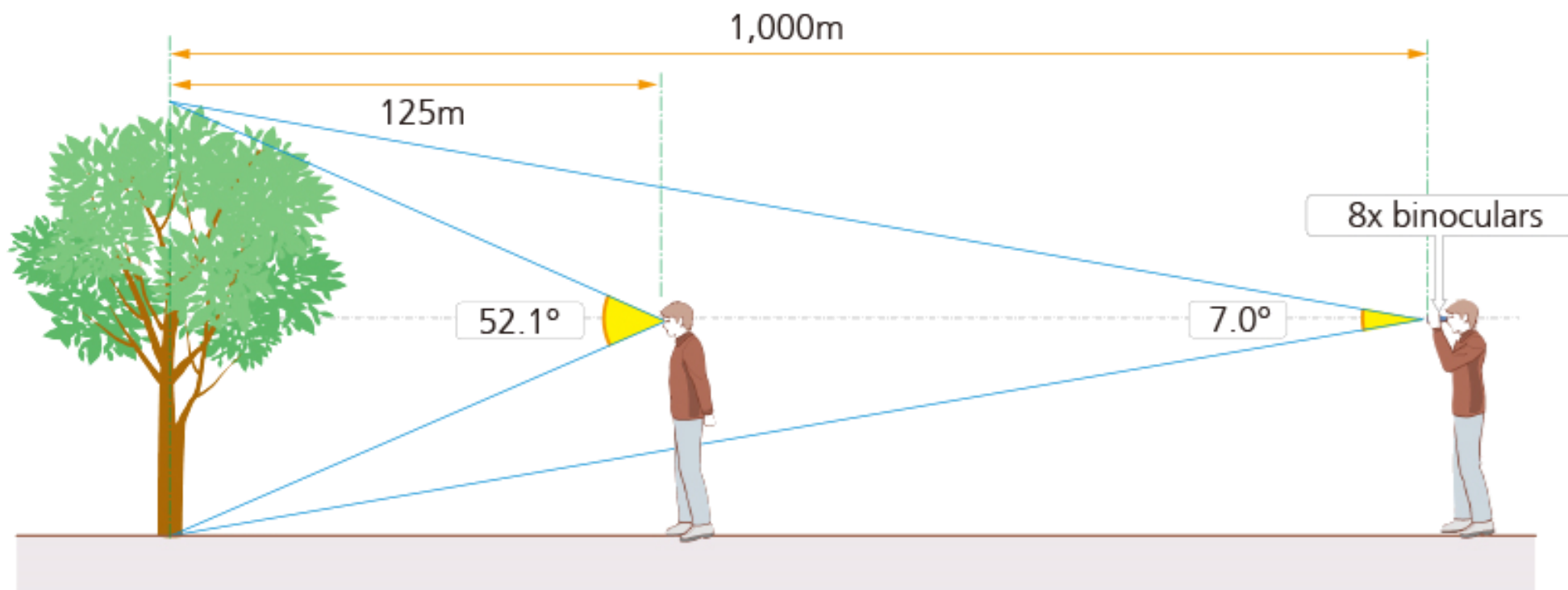
میدان دید (FOV)

- میدان دید برابر با زاویه میدان قابل مشاهده بدون حرکت دوربین است

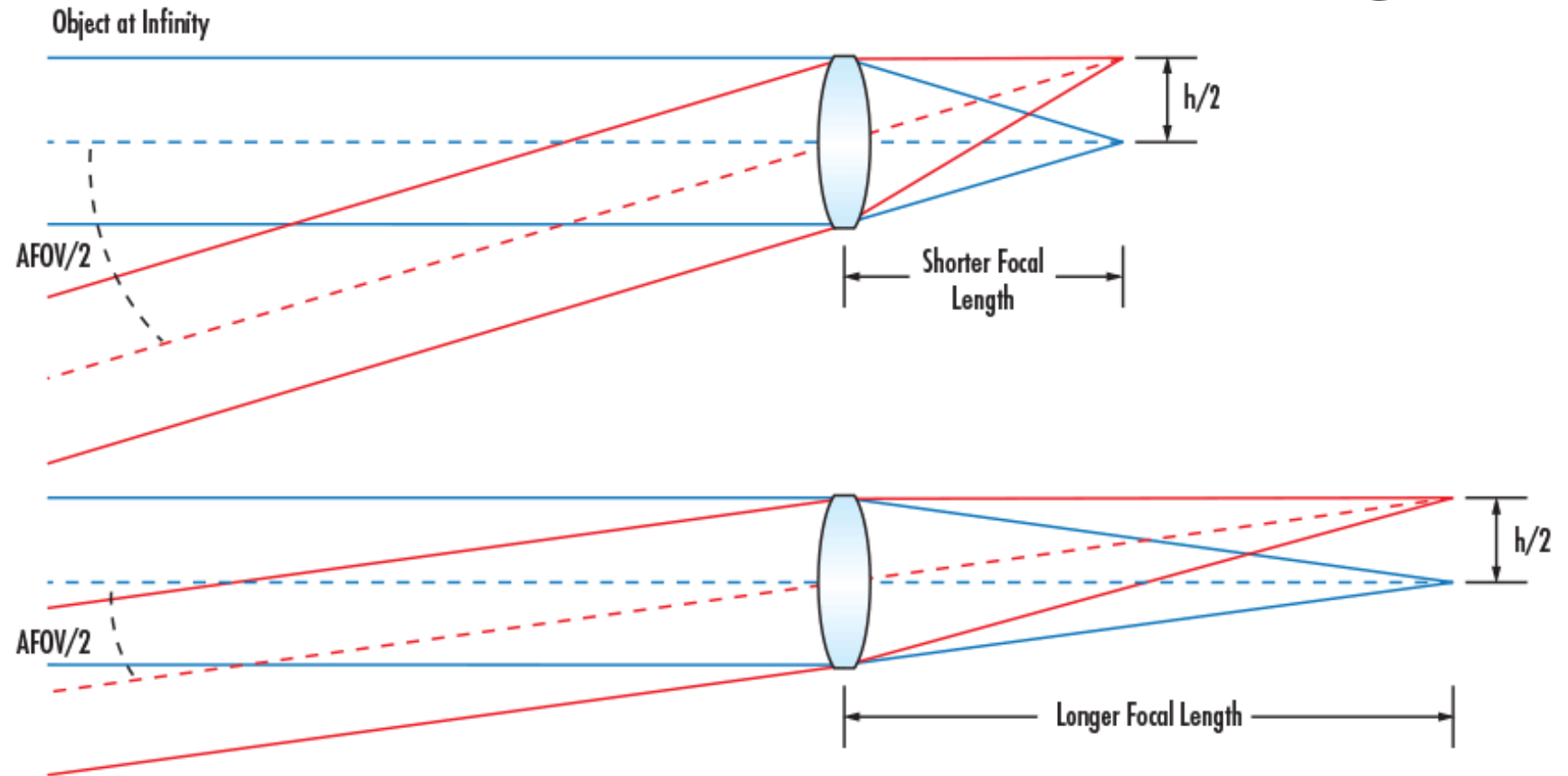
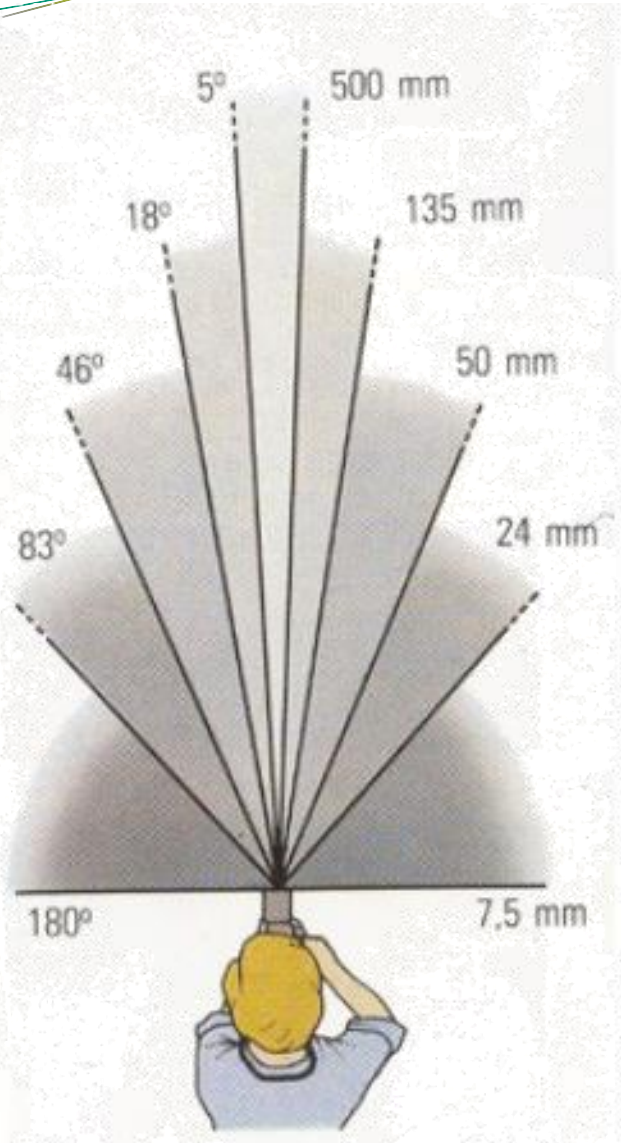


میدان دید (FOV)

- میدان دید برابر با زاویه میدان قابل مشاهده بدون حرکت دوربین است

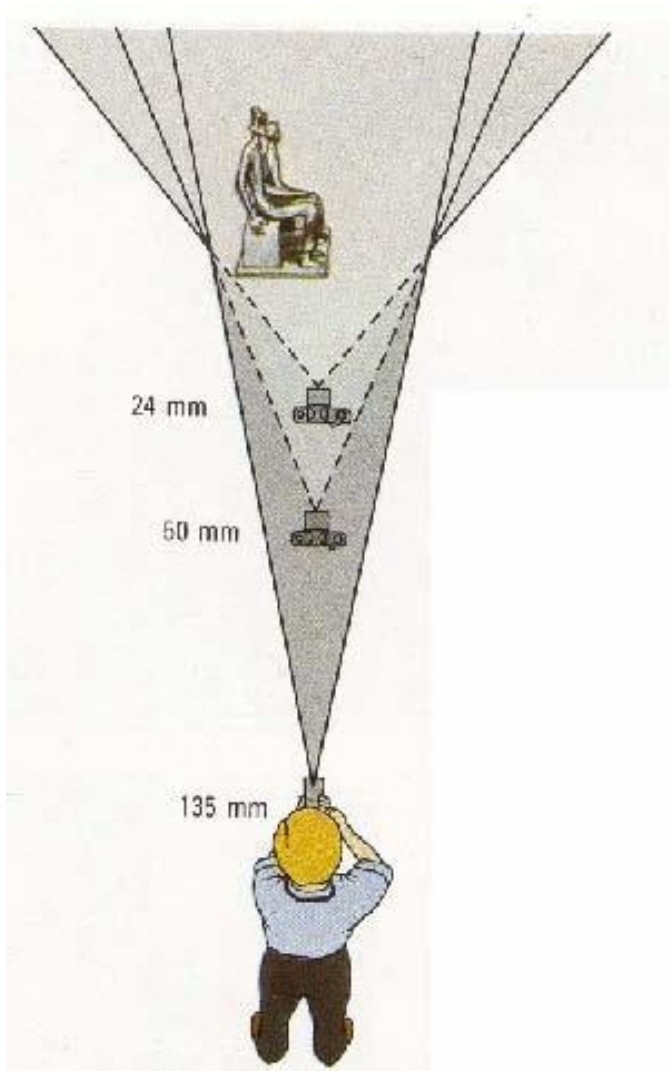


میدان دید (FOV)



$$AFOV = 2 \tan^{-1} \left(\frac{h}{2f} \right)$$

اعوجاج چشم انداز



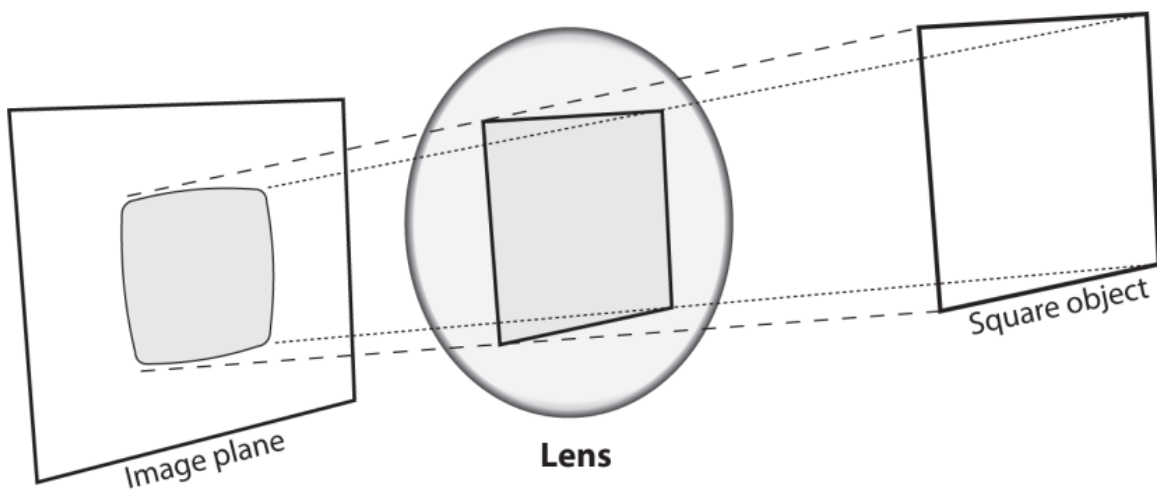
shutterstock.com · 245172952

لنز Fisheye



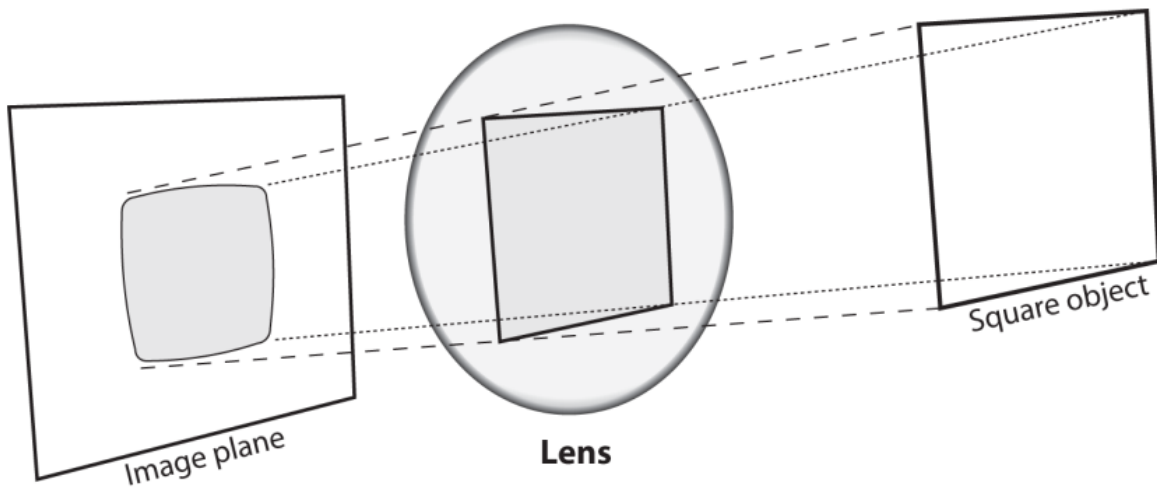
اعوجاج لنز (Lens Distortion)

- در عمل، هیچ لنز ساخته شده‌ای ایده‌آل نیست
- دو نوع اعوجاج اصلی وجود دارد:
 - اعوجاج شعاعی (Radial) که حاصل از شکل لنز است
 - اعوجاج مماسی (Tangential) که حاصل از فرآیند سوار کردن دوربین است



اعوجاج شعاعی

- به خصوص در لنزهای ارزان قیمت این اعوجاج به سادگی قابل تشخیص است
- اعوجاج در مرکز نوری ۰ است و با حرکت به سمت مرز تصویر افزایش می یابد
- این اعوجاج نسبت به مرکز متقارن فرض می شود و با استفاده از چند عبارت از سری تیلور می توان آن را مدل کرد



$$x_{corrected} = x \cdot (1 + k_1 r^2 + k_2 r^3 + k_3 r^6)$$

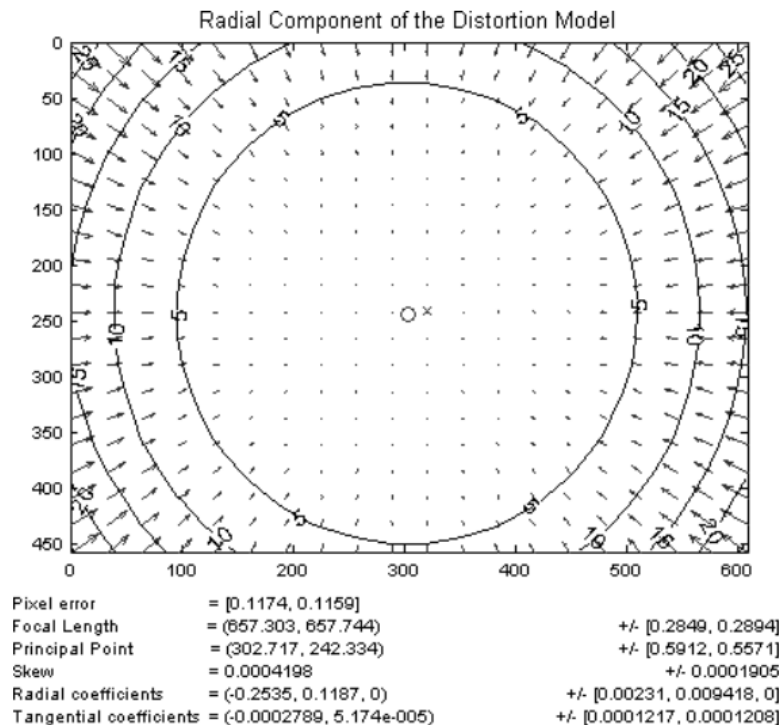
$$y_{corrected} = y \cdot (1 + k_1 r^2 + k_2 r^3 + k_3 r^6)$$

اعوجاج شعاعی

$$x_{corrected} = x \cdot (1 + k_1 r^2 + k_2 r^3 + k_3 r^6)$$

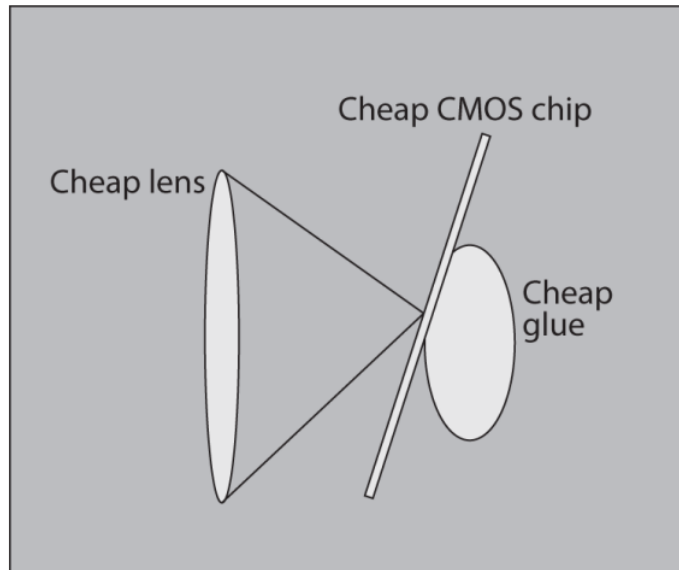
$$y_{corrected} = y \cdot (1 + k_1 r^2 + k_2 r^3 + k_3 r^6)$$

- که (x, y) مختصات اولیه برای نقطه دچار اعوجاج شده است و $(x_{corrected}, y_{corrected})$ مختصات جدید آن بعد از اصلاح است

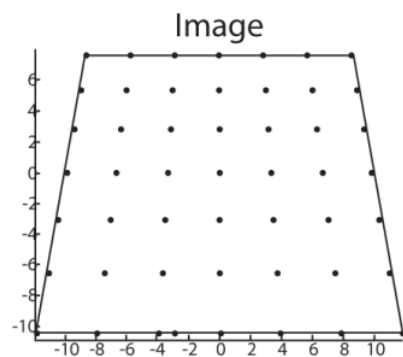


اعوجاج مماسی

- این اعوجاج در فرآیند ساخت دوربین ایجاد می‌شود و به دلیل موازی نبودن دقیق صفحه تصویربرداری با لنز است
- معمولاً به صورت زیر مدل می‌شود



Cheap camera



$$x_{corrected} = x + [2p_1xy + p_2(r^2 + 2x^2)]$$

$$y_{corrected} = y + [p_1(r^2 + 2y^2) + 2p_2xy]$$

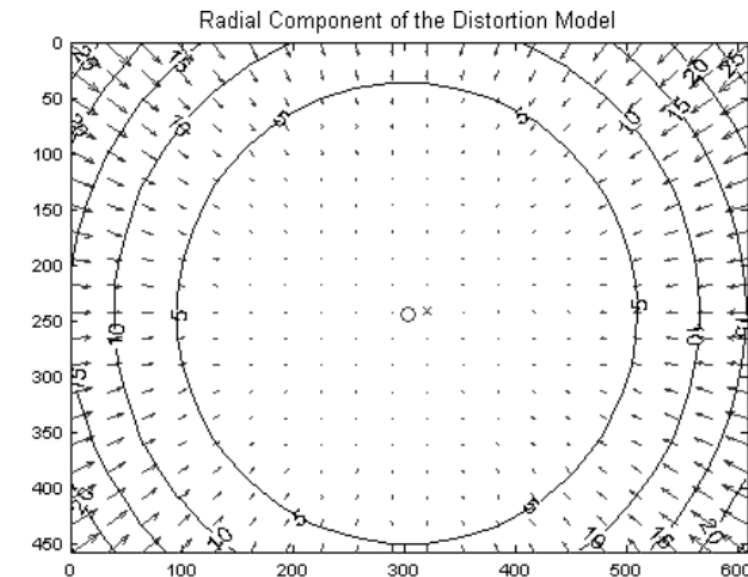
اعوجاج شعاعی و مماسی

$$x_{corrected} = x \cdot (1 + k_1 r^2 + k_2 r^3 + k_3 r^6)$$

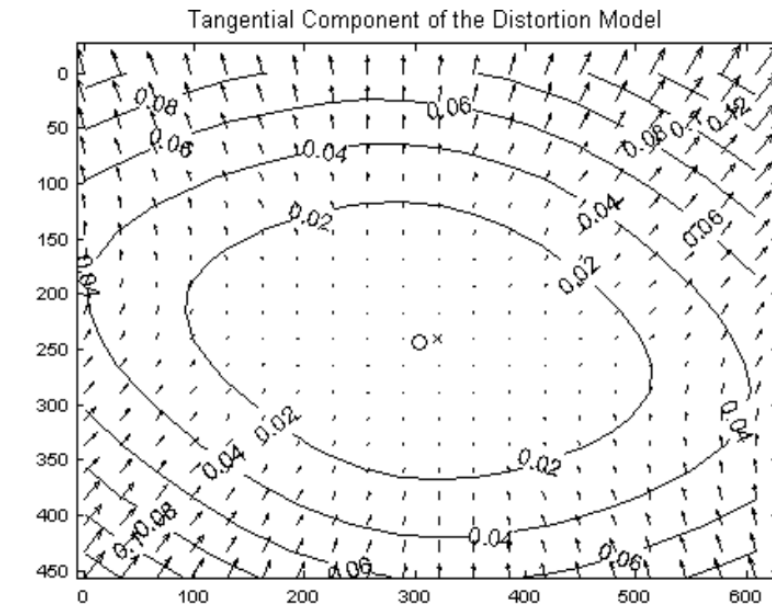
$$x_{corrected} = x + [2p_1 xy + p_2(r^2 + 2x^2)]$$

$$y_{corrected} = y \cdot (1 + k_1 r^2 + k_2 r^3 + k_3 r^6)$$

$$y_{corrected} = y + [p_1(r^2 + 2y^2) + 2p_2 xy]$$



Pixel error = [0.1174, 0.1159]
 Focal Length = (657.303, 657.744) +/- [0.2849, 0.2894]
 Principal Point = (302.717, 242.334) +/- [0.5912, 0.5571]
 Skew = 0.0004198 +/- 0.0001905
 Radial coefficients = (-0.2535, 0.1187, 0) +/- [0.00231, 0.009418, 0]
 Tangential coefficients = (-0.0002789, 5.174e-005) +/- [0.0001217, 0.0001208]



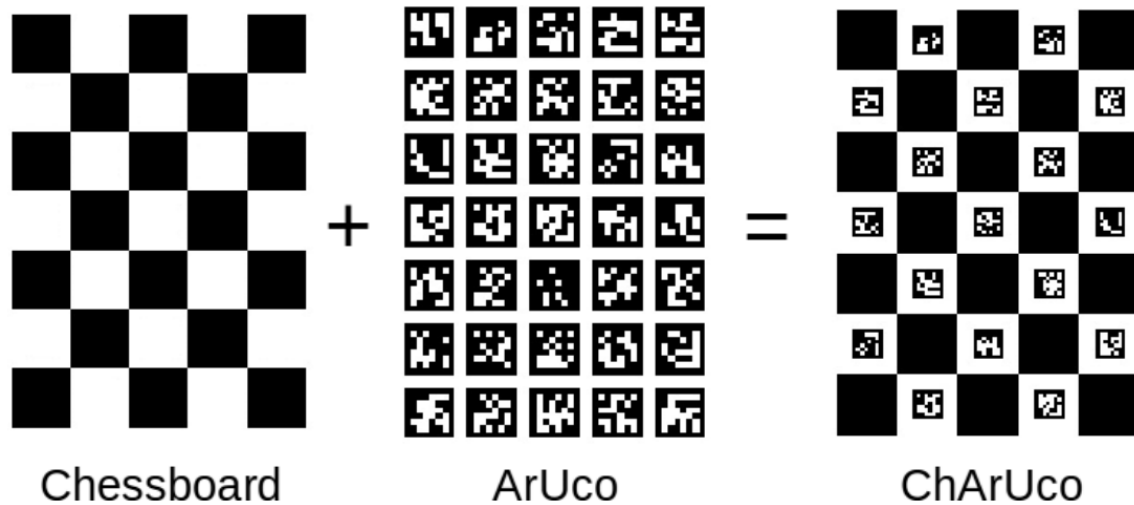
Pixel error = [0.1174, 0.1159]
 Focal Length = (657.303, 657.744) +/- [0.2849, 0.2894]
 Principal Point = (302.717, 242.334) +/- [0.5912, 0.5571]
 Skew = 0.0004198 +/- 0.0001905
 Radial coefficients = (-0.2535, 0.1187, 0) +/- [0.00231, 0.009418, 0]
 Tangential coefficients = (-0.0002789, 5.174e-005) +/- [0.0001217, 0.0001208]

• در مجموع ۵ پارامتر دارند:

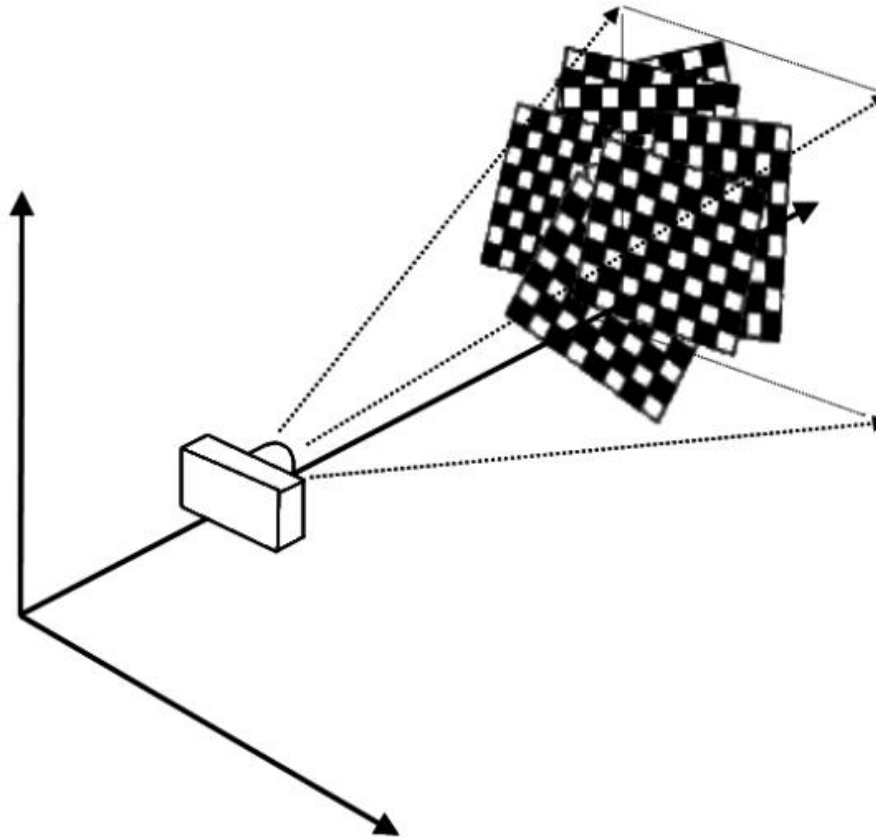
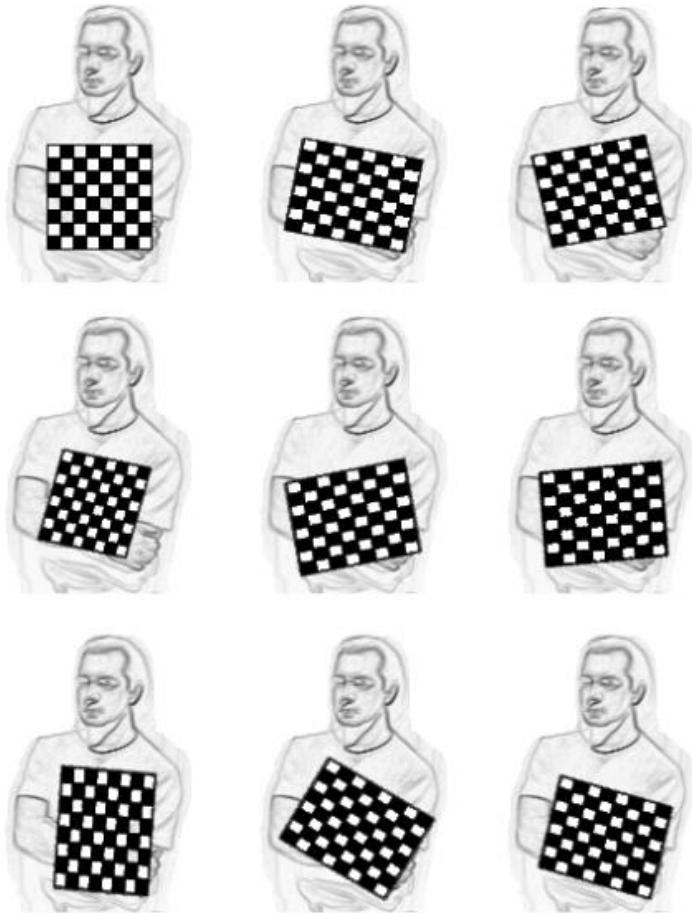
k_3, p_2, p_1, k_2, k_1

Calibration

- می‌خواهیم پارامترهای مربوط به اعوجاج‌های دوربین را تخمین بزنیم
- برای این کار، یک الگوی دارای ساختار مشخص که تعداد زیادی نقطه متمایز و قابل تشخیص داشته باشد را روبروی دوربین قرار می‌دهیم
- با مشاهده این ساختار از زوایای مختلف، می‌توانیم موقعیت نسبی دوربین در بین فریم‌ها و همچنین پارامترهای داخلی (intrinsic) دوربین را محاسبه کنیم



Calibration



طيف الكتر ومغناطيسي

