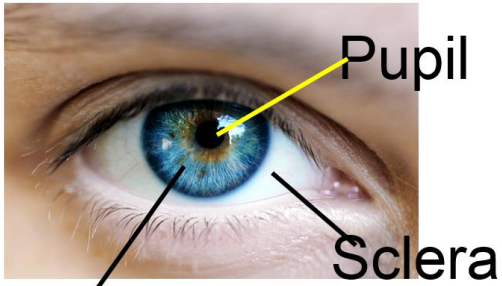
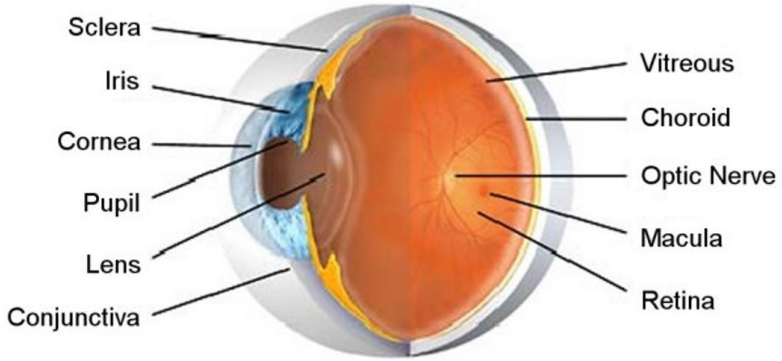
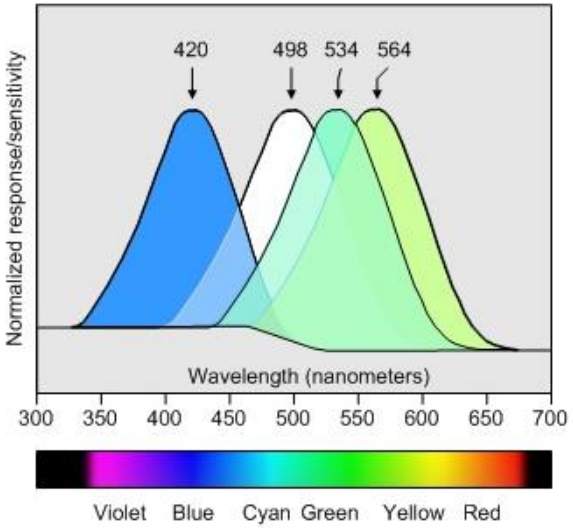
Vision

Human Eye

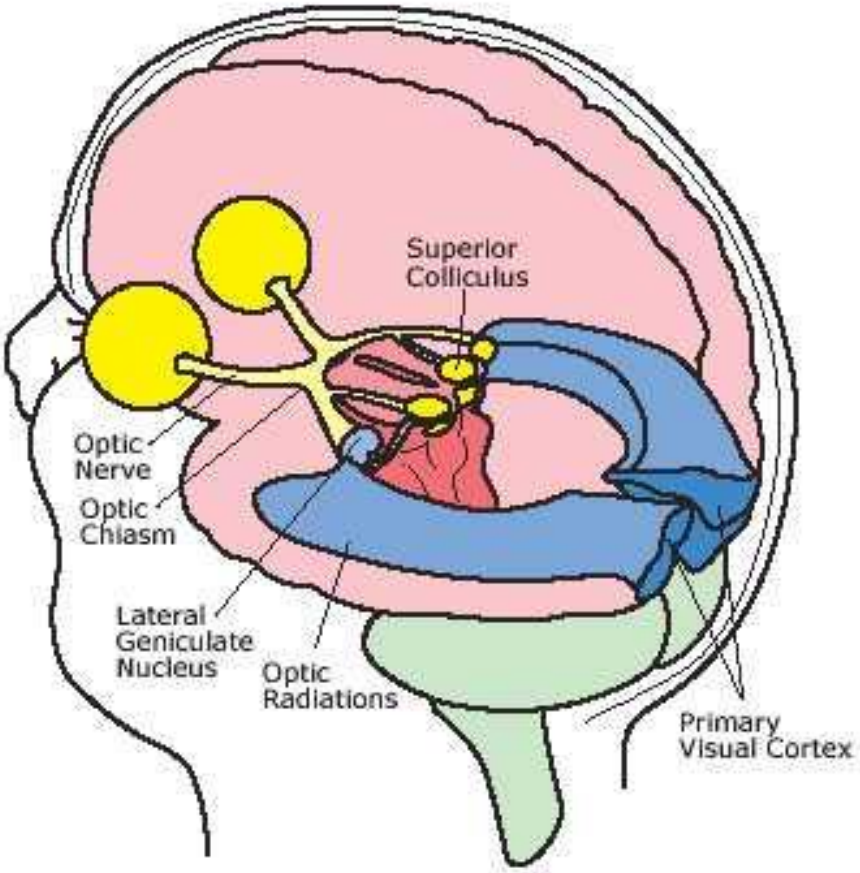






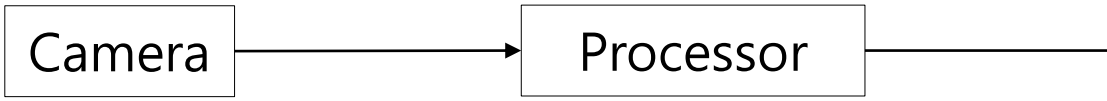
Tristimulus

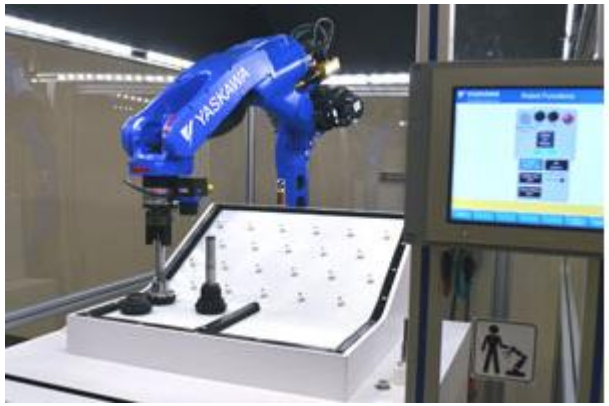
Human Vision System



- human brain
  - 1.5 kg
  - 10^{11} neurons
  - ~1/3 for vision

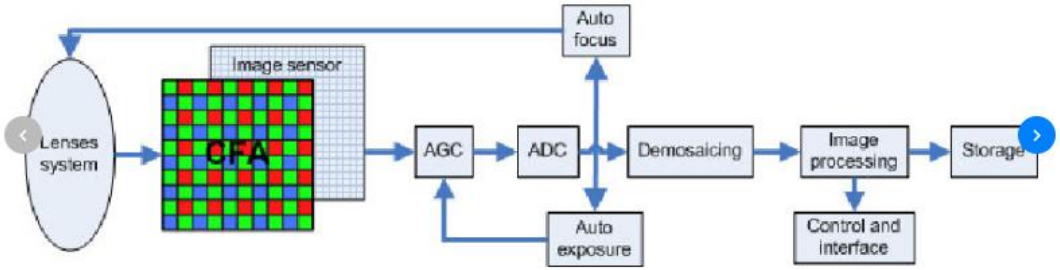
Robot Vision System





Camera

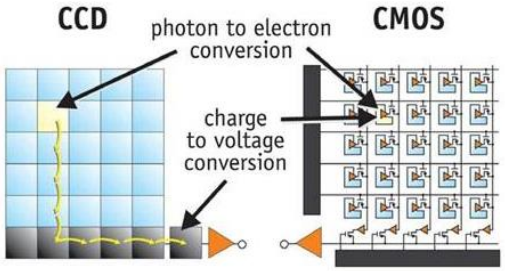
- Camera
  - Image sensor + color filter + lens + ADC + …



- Image Sensor
  - Changes light to electrons

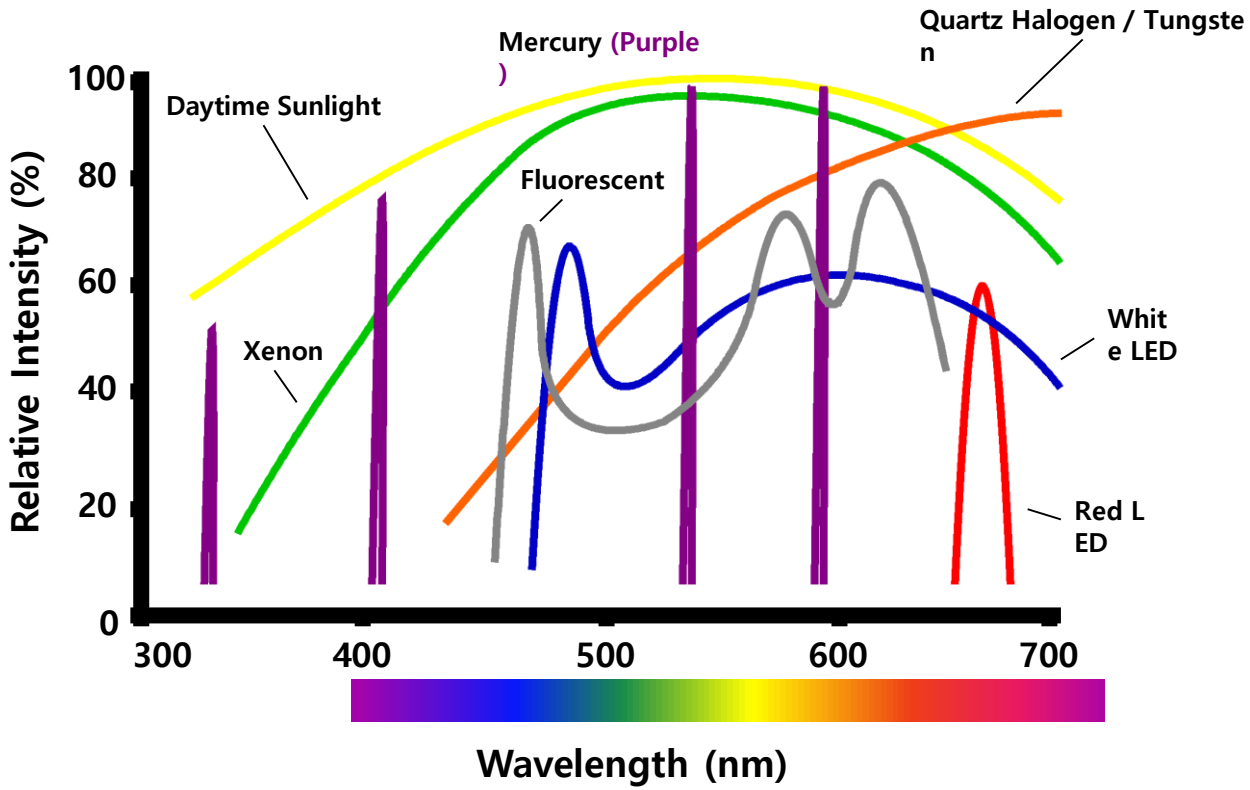
1) CCD
   - charge coupled device
   - High-quality & lower noise image

2) CMOS
   - complementary metal oxide semiconductor
   - Less expensive
   - More noise
   - Need lower power



Lighting

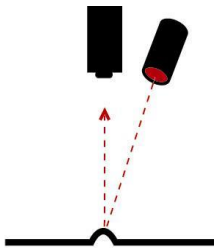
- Lighting source
  - LED (Light Emitting Diode)
  - Quartz Halogen
  - Fluorescent (형광등)
  - Xenon (High-Performance Strobing)



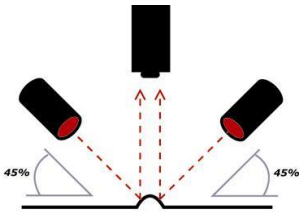
Lighting

- Type
  - Bright field lighting
  - Dark field lighting
  - Defuse lighting
  - Back lighting

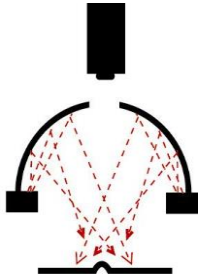
  * Structured lighting



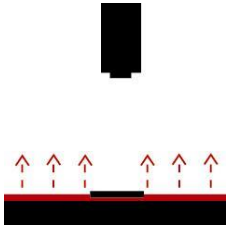
Bright field lighting



Dark field lighting



Diffuse lighting



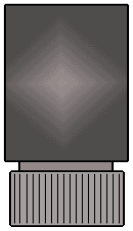
Back lighting

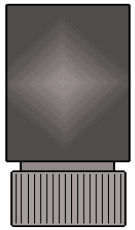
Lighting

- Bright field vs. Dark field

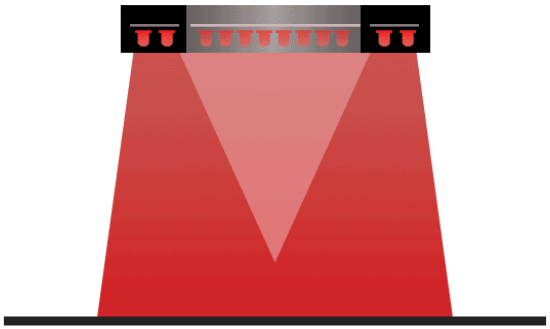
Typical Co-axial Ring Light

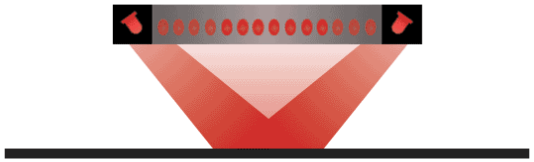
Bright Field





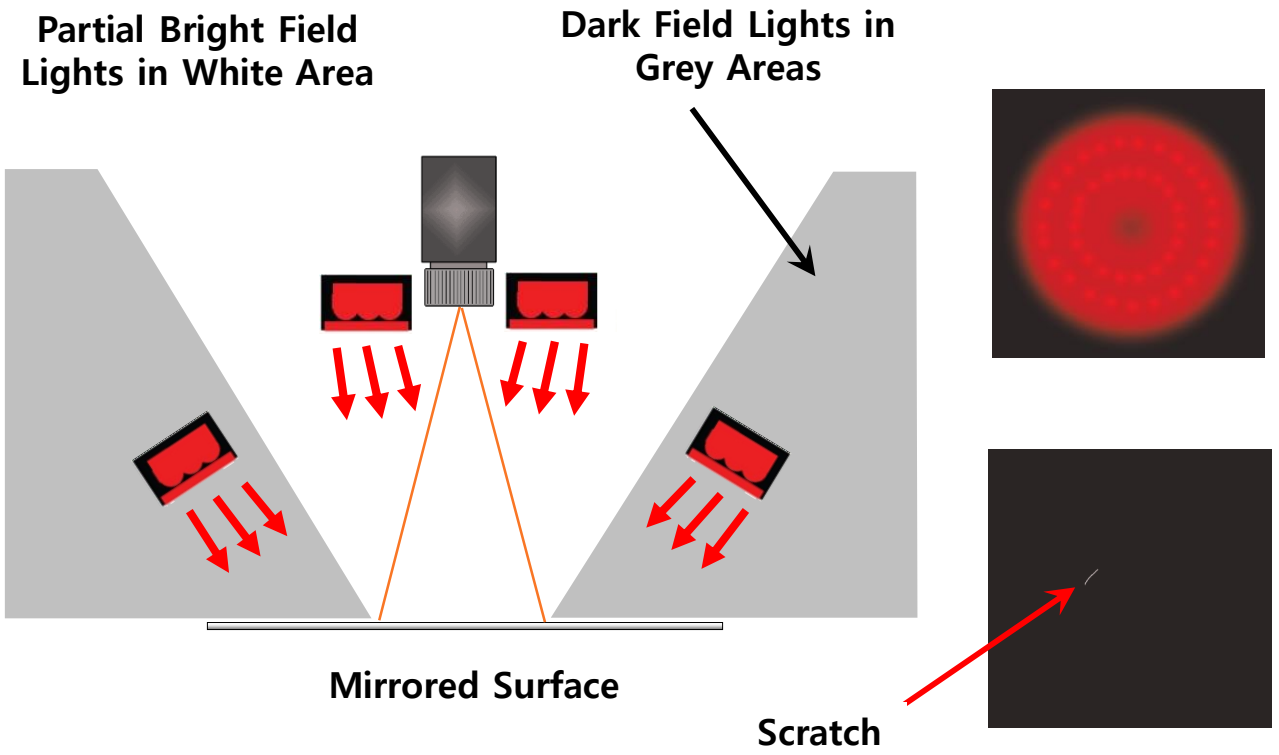
Dark Field





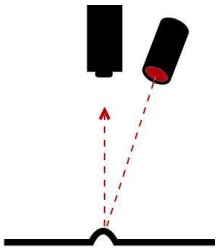
Lighting

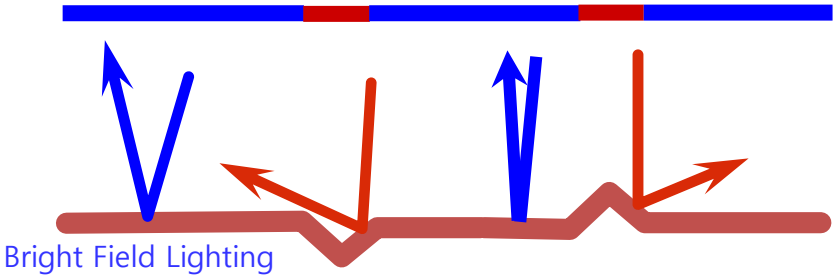
- Bright field vs. Dark field



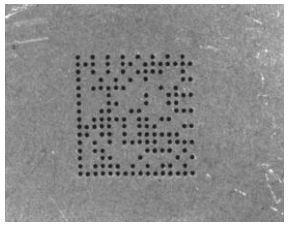
Lighting

- Bright field vs. Dark field

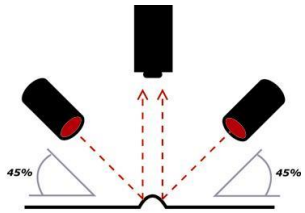


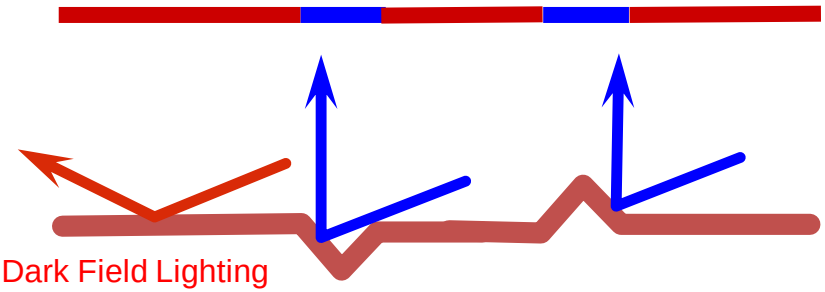


Bright Field Lighting

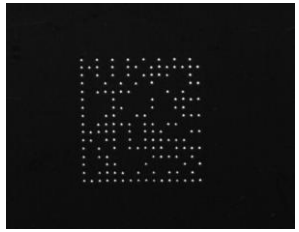


- Similar to sunlight
- Effective for flat surface





Dark Field Lighting

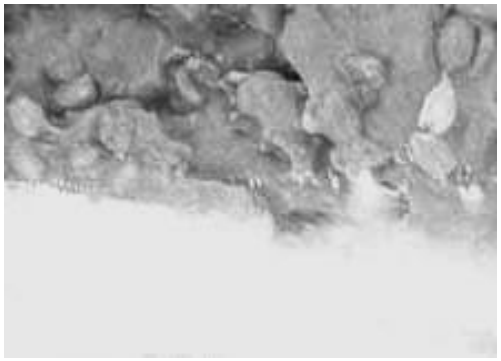


- Effective for specular surface

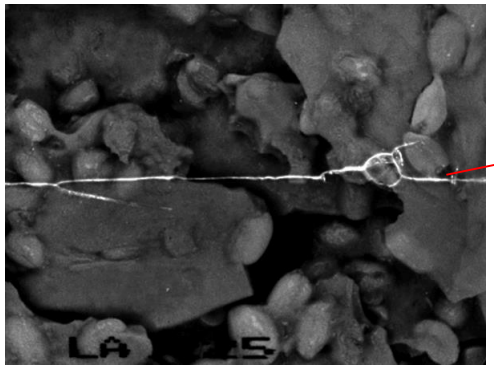
Lighting

- Bright field vs. Dark field

Peanut Brittle Bag



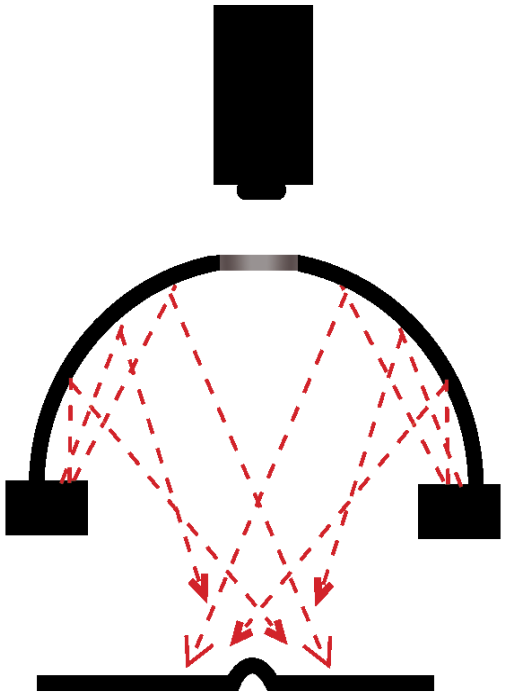
Bright field ring light



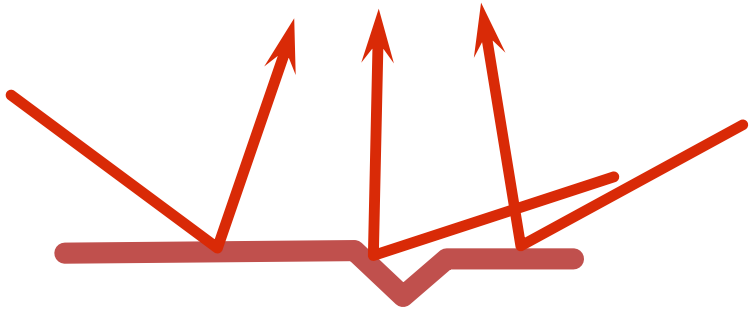
Dark field ring light

Lighting

- Diffuse lighting

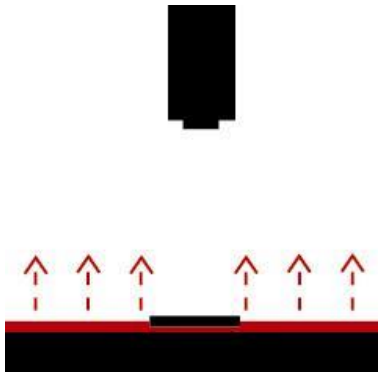


- Similar to the light on an overcast day.
- Creates minimal glare.
  - Surface Texture Is Deemphasized
  - Best Choice for Curved Shiny Parts



Lighting

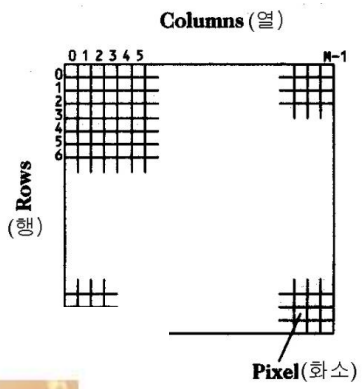
- Back lighting

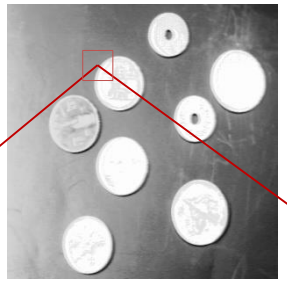


- Creates dark silhouettes against bright background
- Detecting
  - Presence / absence of holes and gaps
  - Part placement or orientation
- Measuring objects
- Useful for translucent materials

Image Data

- 1 Frame = M x N pixels
  - N bits / pixel
    - Gray image: max N = 8
    - Color image: max N = 8 (red) +8 (green) + 8 (blue)





138	138	141	138	138	138	138	138	138	138	131	131	127	127	127
141	141	138	138	138	138	138	138	131	131	138	138	131	131	127
138	138	138	138	141	141	138	138	138	138	131	131	138	138	131
138	138	141	138	138	138	138	138	138	138	138	138	131	131	131
138	138	141	141	141	141	138	138	131	131	127	131	127	138	138
141	141	141	144	141	138	138	138	138	138	138	138	138	138	138
141	141	144	144	141	138	141	138	138	138	138	138	138	138	138
148	148	148	148	148	144	144	144	141	141	141	141	141	141	141
150	150	150	152	154	157	157	159	157	154	152	150	150	150	148
174	191	204	214	223	224	225	225	224	224	223	215	207	195	177
225	227	229	253	253	253	253	253	235	235	235	253	253	235	227
253	254	254	253	253	253	235	235	253	253	235	253	253	253	253
253	229	227	226	226	226	227	227	227	227	228	229	231	235	235
227	227	228	228	229	231	231	235	235	235	235	235	235	235	235
229	231	231	235	235	235	235	235	235	235	235	235	235	235	235

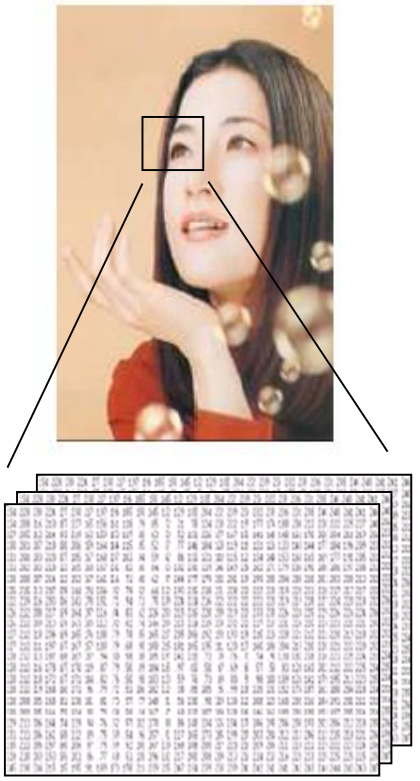


Image Processing

Filtering
Morphology

Segmentation
Line detection
Corner detection

Matching
Learning

Deep Learning

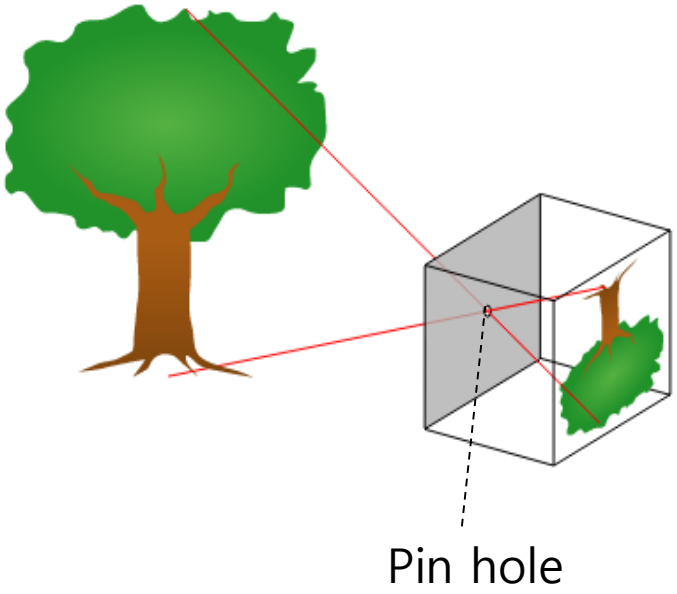
(Low level)

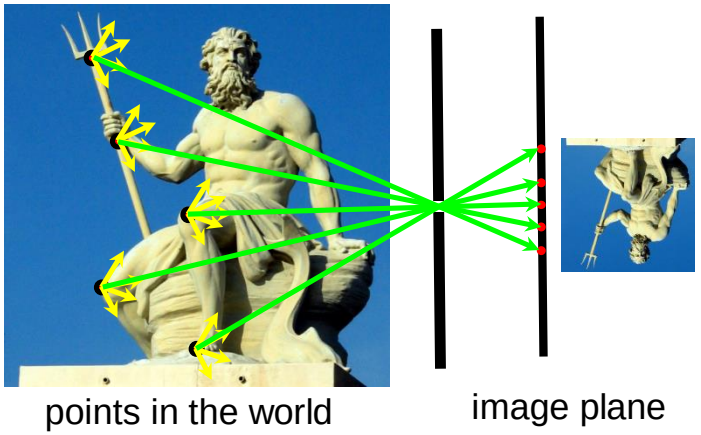
(Medium level)

(High level)

Camera Model

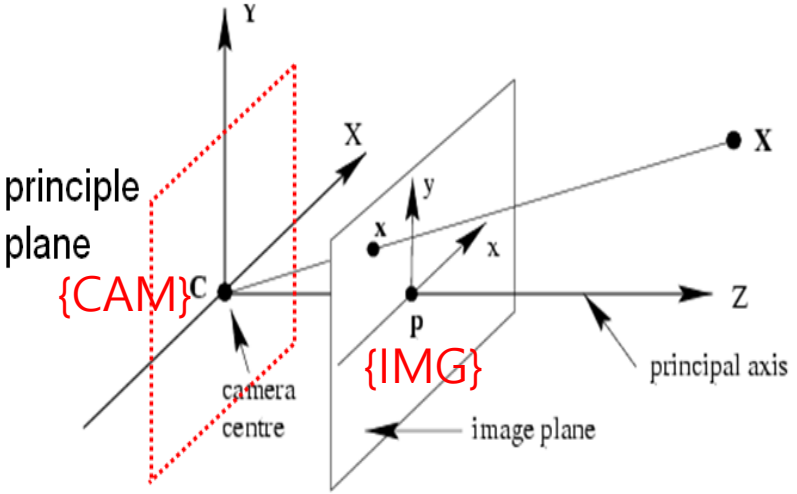
- 핀 홀 (Pin Hole) 모델
  - 카메라를 작은 바늘구멍 (pin hole) 로 모델링
  - 공간 상의 한 점이 작은 바늘 구멍을 통과하여 영상 면에 상이 맺힘



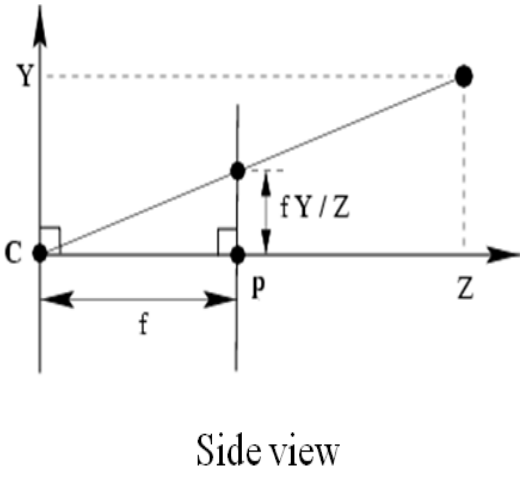


Camera Model

- 카메라 좌표계: {CAM}
  - 카메라 중심을 기준으로 하는 3차원 직교 좌표계
  - C: Projection center / optical center
  - Principal axis (주축)
- 영상 좌표계 {IMG}
  - 영상 평면에서의 2차원 직교 좌표계
  - P : principal point (주축점)



f: focal length



Side view

Camera Model

- Projection

  (X, Y, Z): {CAM} 기준 공간 상 한 점의 좌표

  (x, y): {IMG} 기준 좌표

$$f : x = Z : X \quad \rightarrow \quad x = f\frac{X}{Z}$$

$$f : y = Z : Y \quad \rightarrow \quad y = f\frac{Y}{Z}$$

- Homogeneous coordinate

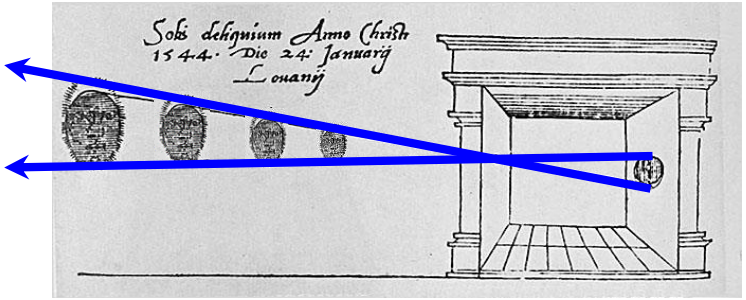
$$\begin{pmatrix} \frac{a}{s} \\ \frac{b}{s} \\ 0 \end{pmatrix} \leftrightarrow \begin{pmatrix} a \\ b \\ s \end{pmatrix}$$

$$\begin{pmatrix} x \\ y \\ 0 \end{pmatrix} = \begin{pmatrix} f\frac{X}{Z} \\ f\frac{Y}{Z} \\ 0 \end{pmatrix} = \begin{pmatrix} fX \\ fY \\ Z \end{pmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

Camera Model

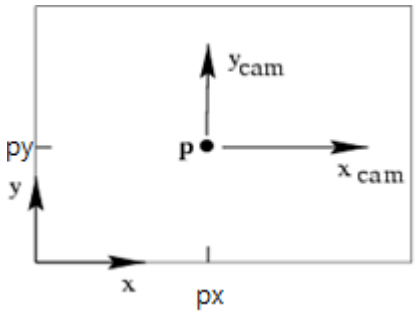
- Perspective Projection
  1) 3D → 2D : $(X, Y, Z) \rightarrow (x, y)$
  2) Straight line → straight line
  3) Parallel lines → lines that intersect at a vanishing point
  4) Conic → conic (circle → circle or ellipse)
  5) Not one-to-one mapping

     X, Y, Z → x, y (OK) , x, y → X, Y, Z (not OK)



Camera Model

- $\boldsymbol{X}_{CAM} \rightarrow \boldsymbol{x}$
  - 영상 면의 중점 편차 고려



$$\begin{pmatrix} x \\ y \\ 0 \end{pmatrix} = \begin{pmatrix} f\dfrac{X}{Z} + p_x \\ f\dfrac{Y}{Z} + p_y \\ 0 \end{pmatrix} = \begin{pmatrix} fX + Z\,p_x \\ fY + Z\,p_y \\ Z \end{pmatrix} = \begin{bmatrix} f & 0 & p_x & 0 \\ 0 & f & p_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

$$= \begin{pmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{pmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

$$\boldsymbol{x} = \boldsymbol{K}\,[\,\boldsymbol{I}\mid\boldsymbol{0}]\,\boldsymbol{X}_{CAM}$$

K: Camera calibration matrix

f, p_x, p_y : Intrinsic parameters (미지수 3 개)

Camera Model

- CCD 카메라
  - 단위계 변환 필요: mm → pixel

$$\boldsymbol{K} = \begin{pmatrix} m_x & 0 & 0 \\ 0 & m_y & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} f & 0 & p_x \\ 0 & f & p_y \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \alpha_x & 0 & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{pmatrix}$$

m_x, m_y: mm 당 pixel 수

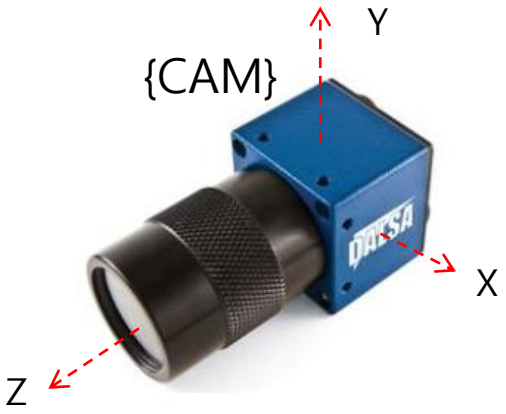
  - 영상 면의 x 축, y 축이 수직하지 않은 경우 고려

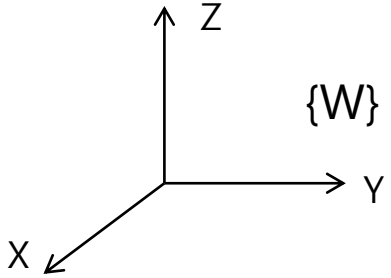
$$\boldsymbol{K} = \begin{pmatrix} \alpha_x & \gamma & x_0 \\ 0 & \alpha_y & y_0 \\ 0 & 0 & 1 \end{pmatrix}$$

미지수 5 개 (Intrinsic parameter)

Camera Model

- 월드 좌표계 (World coordinate) {W}
  - 위치 측정의 기준이 되는 좌표계
  - 카메라의 회전 및 이동 고려





P ●

$$X_{WORLD} = {}^{W}T_{CAM}\, X_{CAM}$$

$${}^{W}T_{CAM} = \begin{bmatrix} R^T & C \\ 0 & 1 \end{bmatrix}$$

R: 3 x 3 rotation matrix
C: 3 x 1 translation vector

Camera Model

- $X_{WORLD} \rightarrow X_{CAM}$

$$X_{CAM} = {}^{W}T_{CAM}{}^{-1} X_{WORLD}$$

$$= \begin{bmatrix} \boldsymbol{R} & -\boldsymbol{R}^T\boldsymbol{C} \\ 0 & 1 \end{bmatrix} X_{WORLD}$$

- $X_{WORLD} \rightarrow x$

$$x = \boldsymbol{K}\,[\,\boldsymbol{I} \mid \boldsymbol{0}]\, X_{CAM}$$
$$= \boldsymbol{K}\,[\,\boldsymbol{I} \mid \boldsymbol{0}] \begin{bmatrix} \boldsymbol{R} & -\boldsymbol{R}^T\boldsymbol{C} \\ 0 & 1 \end{bmatrix} X_{WORLD}$$
$$= \boldsymbol{K}\,[\,\boldsymbol{R} \mid -\boldsymbol{R}^T\boldsymbol{C}]\, X_{WORLD}$$
$$= \boldsymbol{K}\,[\,\boldsymbol{R} \mid \boldsymbol{t}]\, X_{WORLD}$$
$$= \boldsymbol{P}\, X_{WORLD}$$

$\boldsymbol{R}$: rotation matrix (3 x 3)
$\boldsymbol{t}$: translation matrix (3 x 1)
⇒ Extrinsic parameters
⇒ 미지수: 6개 (회전3, 평행3)

$\boldsymbol{P}$: Projection matrix

Camera Calibration

- 목적
  - 카메라를 사용한 3차원 계측을 위하여, 영상 좌표값으로 부터 월드 좌표값을 계산하여야 함
  - 이때 Projection matrix 의 원소값이 얼마나 정확한 지에 따라 계산의 정확도가 달라짐

$$x = \boldsymbol{P}\, X_{WORLD}$$

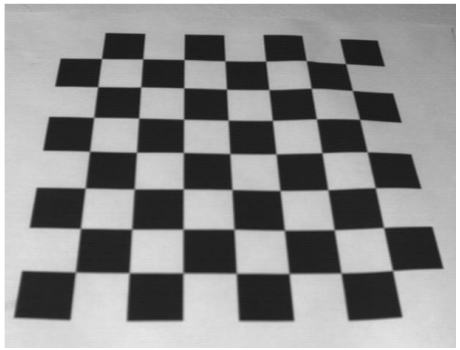
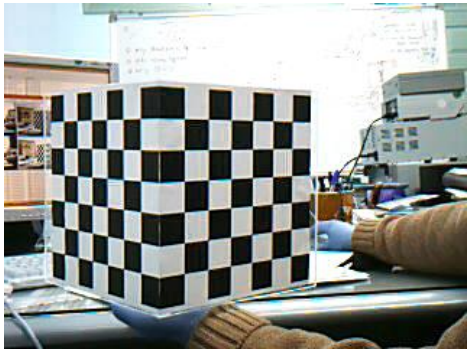
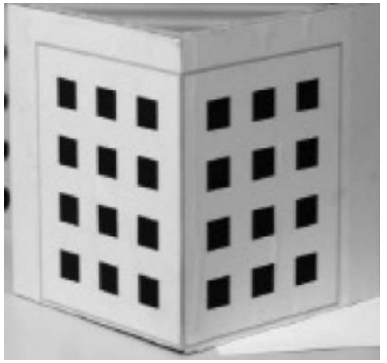
  - Projection matrix 의 원소 값을 결정하는 과정을 카메라 보정 (camera calibration) 이라고 함

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

내부변수 5 개
외부변수 6 개

Camera Calibration

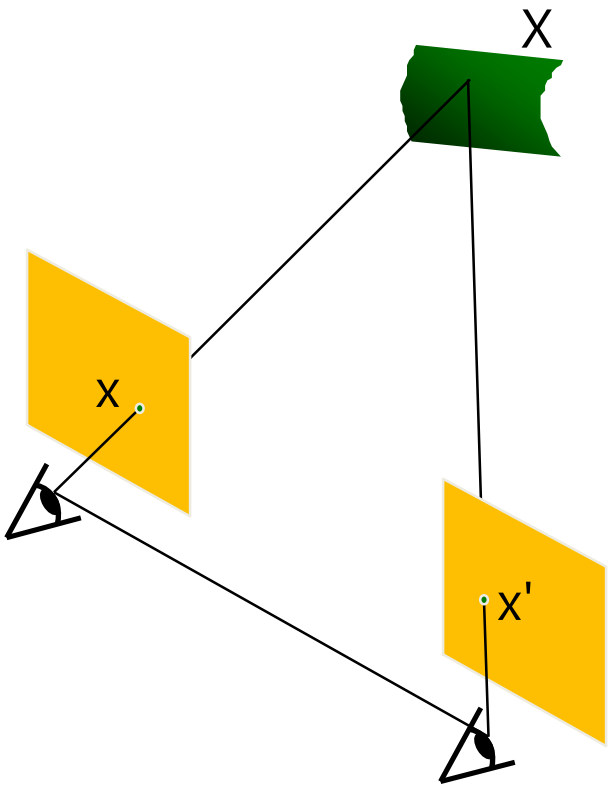
- 방법
  - 치수가 미리 알려져 있는 정밀한 보정자 (calibration scale) 사용

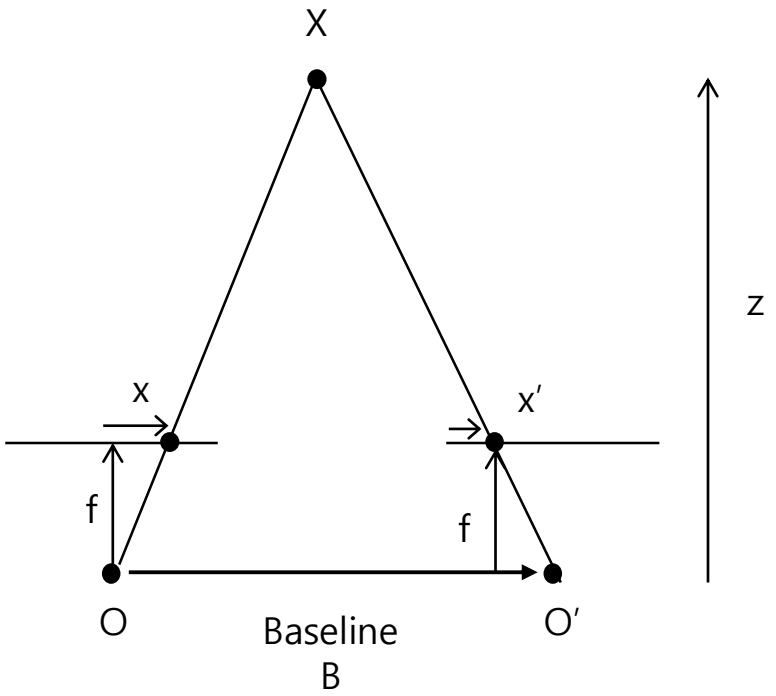
카메라 보정용 스케일의 예

Stereo Vision

- Recover depth by finding image coordinate x′ that corr esponds to x



$$\frac{x - x'}{O - O'} = \frac{f}{z}$$



$$disparity = x - x' = \frac{B \cdot f}{z}$$

Disparity is inversely proportional to depth.

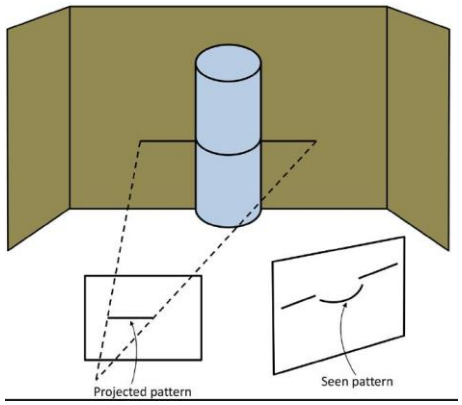
Stereo Vision

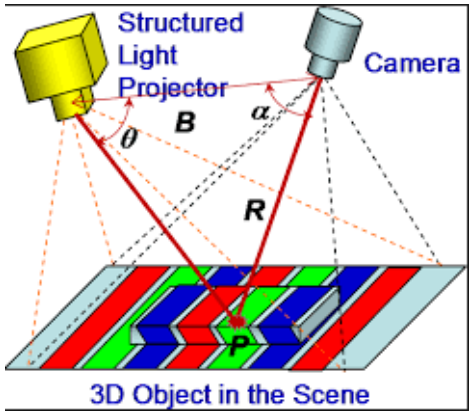
- Correspondence Problem

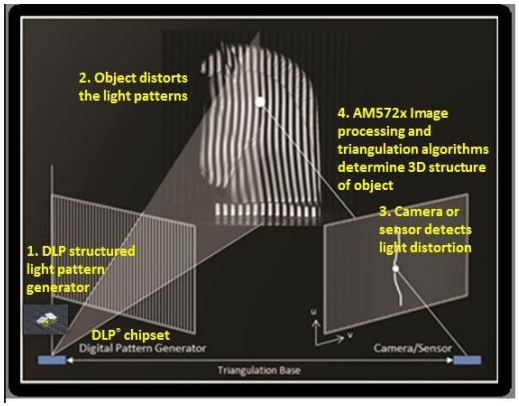
  - We have two images taken from cameras with different i ntrinsic and extrinsic parameters
  - How do we match a point in the first image to a point i n the second? How can we constrain our search?

Structured Lighting

- Recover depth from pattern by structured lighting
  - Projector + Camera
  - Restore depth data from image







Kinect sensor