

Assignment 1 Vision I

Mario Alberto Tapia

de la Cruz

1.-

a) What are approximate values of the focal distances of these cameras in mm?

Note: the (d) means that the field of view is measured along the sensor diagonal.

Tenemos la sig. info

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

$$\tan \left(\frac{\beta}{2} \right) = \frac{d}{2f} \Rightarrow f = \frac{d}{2 \tan \left(\frac{\beta}{2} \right)}$$

Navcam

$$F.o.V : 166.5^\circ \pm 3^\circ(d)$$

RTE

$$102^\circ(h) \times 82^\circ(v), 125^\circ(d)$$

$$\text{Sensor size: } 1.968 \text{ mm} \times 1.488 \text{ mm}$$

$$6.1 \text{ mm} \times 4.574 \text{ mm}$$

$$\text{Lens size: } 1/2.5'' (\sim 2.3 \text{ mm diagonal})$$

$$1/3.06'' (\sim 5.6 \text{ mm diagonal})$$

De manera que para navcam

$$f = \frac{d}{2 \tan \left(\frac{\theta}{2} \right)} = \frac{2.33}{2 \tan \left(\frac{166.5}{2} \right)} = 0.1378 \text{ mm}$$

y para RTE

$$f = \frac{d}{2 \tan \left(\frac{\theta}{2} \right)} = \frac{5.6}{2 \tan \left(\frac{125}{2} \right)} = 1.4575 \text{ mm}$$

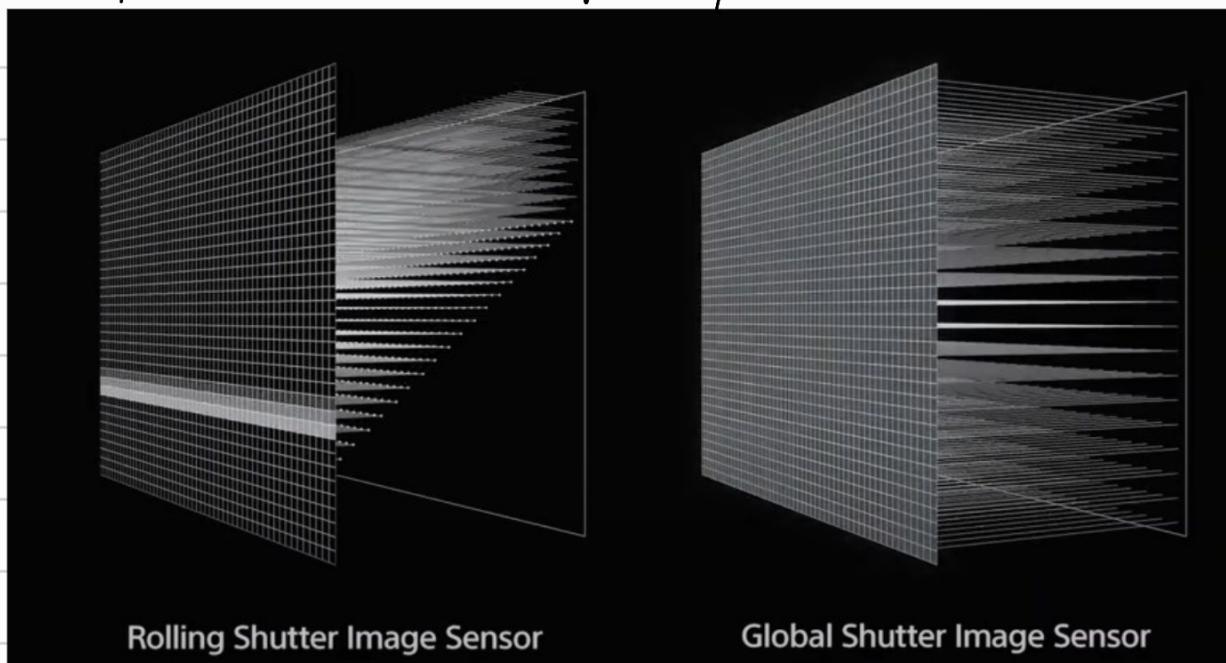
b) How does the RTE camera output color images?

- Tiene un formato de escala de pixel de 0.53 msrad/pixel
con un formato de pixel de color RGB (8 bit Bayer)

c) Research what are the differences between global shutter (in the Navcam camera) vs. rolling shutter (in the RTE camera)

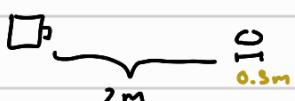
Global Shutter: Hace la exposición y lee todos los pixeles de manera simultánea

Rolling Shutter: Hace la exposición y lee los pixeles de forma fila por fila, por lo que se pueden presentar deformaciones en la foto final de cosas que se muevan más rápido que la velocidad del obturador.



d) If the helicopter takes photos of a 0.5m-wide rock located 2m in front of the camera, and centered on its optical axis, what should be the lens-CCD distance for the rock to be well focused?

$$\text{Ahora usaremos } \frac{1}{f} = \frac{1}{d_1} + \frac{1}{d_2}$$



de manera que f es el calculado y

$d_1 = 2\text{m} = 2000\text{mm}$, hay que calcular d_2 , despejando,

tenemos

$$\frac{1}{f} - \frac{1}{d_1} = \frac{1}{d_2} \Rightarrow d_2 = \frac{1}{\frac{1}{f} - \frac{1}{d_1}}$$



Para RTE

$$d_2 = \frac{\frac{1}{1.4575} - \frac{1}{2000}}{\frac{1}{1.4575} - \frac{1}{2000}} = 1.4585\text{ mm}$$

Para navcam

$$d_2 = \frac{\frac{1}{0.1378} - \frac{1}{2000}}{\frac{1}{0.1378} - \frac{1}{2000}} = 0.1378\text{ mm}$$

c) In that case, what would be the size of the image of the rock in pixels?

Ahora el tamaño de la imagen de la roca en pixeles

Tenemos la siguiente información

	Navcam	RTE
Tamaño de pixel	$3\mu\text{m} \times 3\mu\text{m} \rightarrow 0.003\text{mm/pixel}$	$1.17\mu\text{m} \times 1.17\mu\text{m} \rightarrow 0.00112\text{mm/pixel}$
d_2	0.1378mm	1.4585mm
d_1	2000mm	2000mm
tamaño del objeto	500mm	500mm

$$1\text{m} \rightarrow \text{micromilimetros} \quad 1\text{m} = 0.001\text{mm}$$

Queremos el tamaño de la imagen en el sensor $M = \frac{d_2}{d_1}$
donde M es la magnificación.

Para navcam

$$M = \frac{d_2}{d_1} = \frac{0.1378\text{mm}}{2000\text{mm}}$$

$$= 6.89 \times 10^{-5}$$

Para RTE

$$M_{\text{RTE}} = \frac{d_2}{d_1} = \frac{1.4585}{2000}$$

$$= 7.2925 \times 10^{-4}$$

Luego el tamaño de la imagen en el sensor

$$\text{Tamaño en mm} = M \times \text{Tamaño del objeto}$$

Para navcam

$$\text{Tamaño} = 6.89 \times 10^{-5} (500\text{mm})$$

$$= 0.03445$$

Para RTE

$$\text{Tamaño} = 7.2925 \times 10^{-4} (500)$$

$$= 0.364625\text{mm}$$

Luego la conversión a pixeles

$$\text{Tamaño en pixeles} = \frac{\text{Tamaño imagen en mm}}{\text{Tamaño del pixel en mm}}$$

Navcam

$$\text{Tamaño pixels} = \frac{0.03445 \text{ mm}}{0.003 \text{ mm/pixel}} = 11.48 \text{ pixels}$$



RTE

$$\text{Tamaño pixels} = \frac{0.364625 \text{ mm}}{0.00112 \text{ mm/pixel}} = 325.5580 \text{ pixels}$$

2. [2pts] In an average human being, the distance between the crystalline (the lens) and the eye fundus is about 16mm. The average refraction index of the crystalline is 1.42. Deduce the range of variation of the radius of the crystalline when we accomodate our eye to focus on scenes between 5cm and ∞ . Plot in Python the corresponding function that gives this radius in function of the depth to be focused on.

Datos:

Distancia cristalino y fundus \rightarrow 16mm

Refraction index of crystalline \rightarrow 1.42

Rango de variación del radio del cristalino al acomodar el ojo para enfocar entre 5cm y ∞

$$\text{Sea } f = \frac{R}{2(n_{\text{glass}} - 1)}, \text{ luego tenemos que } \frac{1}{d_1} + \frac{1}{d_2} = \frac{1}{f}$$

donde

d_1, d_2 distancias desde la intersección del rayo con el eje óptico al centro óptico

Veamos los siguientes casos

$$n_{\text{glass}} = 1.42$$

$$d_2 = 16\text{mm}$$

$$\frac{1}{f} = \frac{1}{d_1} + \frac{1}{d_2} = \frac{1}{d_2}$$

$$\text{Por lo que } \frac{1}{f} = \frac{1}{d_2} \Rightarrow f = d_2 = 16\text{mm}$$

Caso 2: Enfocando a 5cm (50mm)

$$\frac{1}{f} = \frac{1}{d_1} + \frac{1}{d_2} \Rightarrow \frac{1}{f} = \frac{1}{50} + \frac{1}{16} = 0.0825$$

$$\Rightarrow f_s = \frac{1}{0.0825} \approx 12.1212$$

ya que hemos encontrado f , procedemos encontrar el radio de curvatura

$$\text{Para } f_\infty = \frac{R}{2(n_{\text{glass}} - 1)}$$

$$\Rightarrow 16 \text{ mm} = \frac{R}{2(1.42 - 1)} \Rightarrow R = 16 (2(1.42 - 1)) \\ R = 13.44 \text{ mm}$$

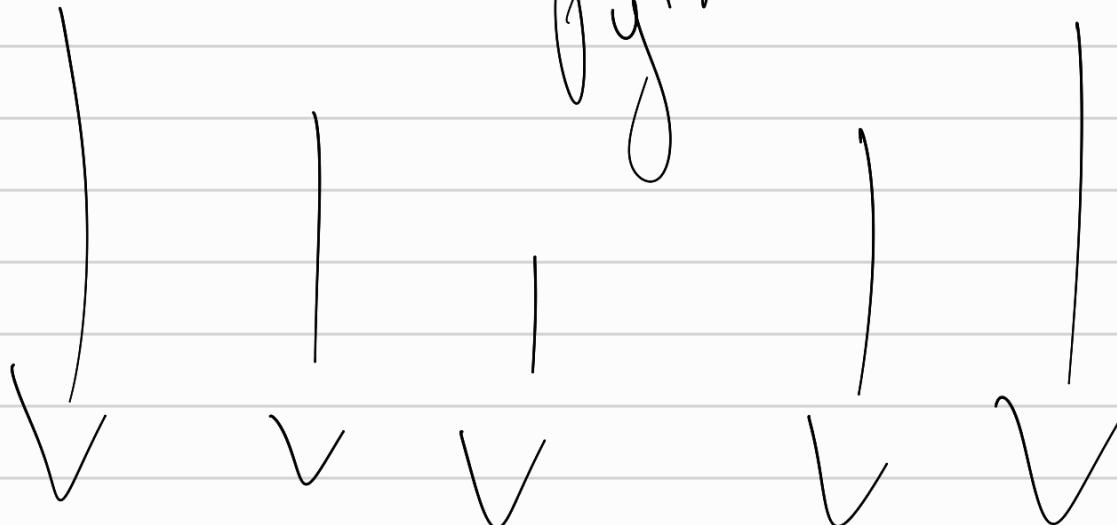
$$\text{Para } f_5 = \frac{R}{2(n_{\text{glass}} - 1)}$$

$$\Rightarrow R = (12.1212)(2(1.42 - 1)) \approx 10.181808$$

Ahora hay que ploteando en Python

En la siguiente página está el código de

Python.



```

import numpy as np
import matplotlib.pyplot as plt

#vamos a calcular el intervalo del rango de variacion del cristalino
#del ojo humano
#tenemos los datos
n_glass = 1.42
d2 = 16.0 #milimetros, esta es la distancia entre el cristalina y el
eye fundus (Retina)
#queremos calcular al acomodar el ojo entre 5 cm (50 mm) y el infinito
d1_5cm = 50.0 #milimetros
d1_inf = np.inf #milimetros

#primero queremos f
def f (d1, d2):
    return 1/(1/d1 + 1/d2)

#luego la función para calcular R
def R (n_glass, f):
    return f*(2*(n_glass - 1))

#ahora calculamos el rango de variación
R_5cm = R(n_glass, f(d1_5cm, d2))
R_inf = R(n_glass, f(d1_inf, d2))
print("El rango de variación del cristalino del ojo humano al acomodar
el ojo entre 5 cm y el infinito es de {:.2f} a {:.2f}
mm".format(R_5cm, R_inf))

```

El rango de variación del cristalino del ojo humano al acomodar el ojo entre 5 cm y el infinito es de 10.18 a 13.44 mm

```

#ahora haremos el plot de la variación del radio
d1_valores = np.linspace(50, 1000, 1000)
R_valores = R(n_glass, f(d1_valores, d2))

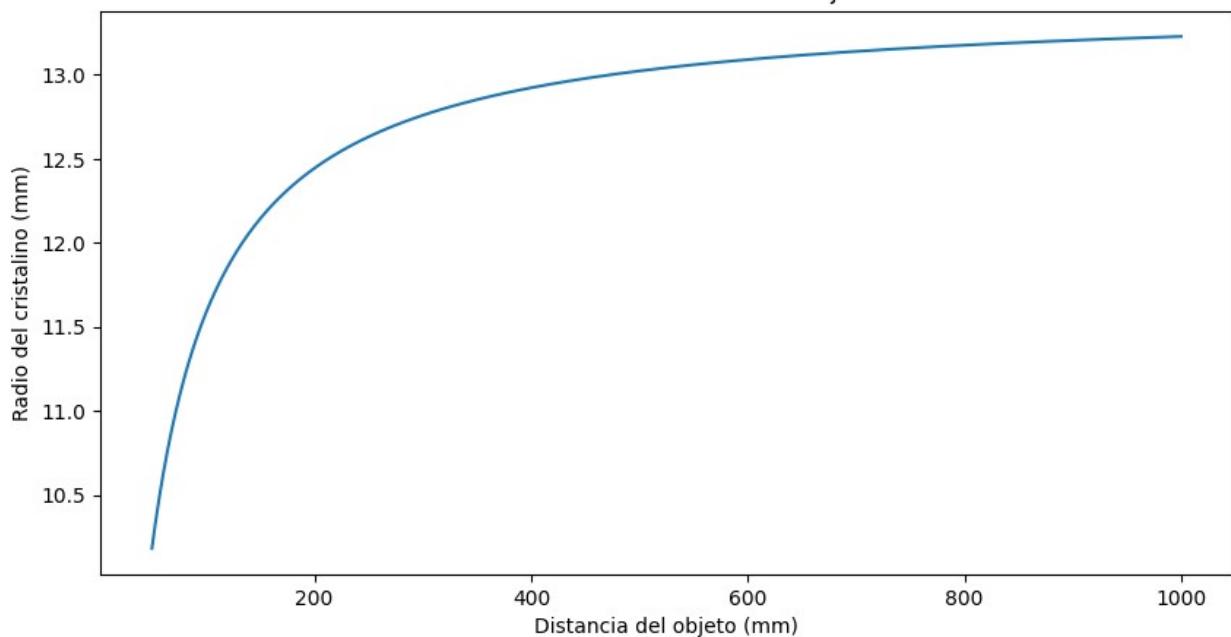
#luego la gráfica
plt.figure(figsize=(10, 5))

plt.plot(d1_valores, R_valores)
plt.xlabel("Distancia del objeto (mm)")
plt.ylabel("Radio del cristalino (mm)")
plt.title("Variación del radio del cristalino del ojo humano")

```

Text(0.5, 1.0, 'Variación del radio del cristalino del ojo humano')

Variación del radio del cristalino del ojo humano



Aquí
termina
el plot :)



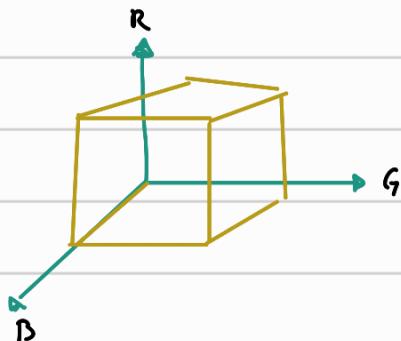
3. [2 pts] It is estimated that the human being can distinguish among around 10 millions colors. Give an approximate value of the number of colors it could distinguish if it were equipped with **4 kinds of cones instead of 3?** (assuming that the sensibility is uniform along cells and wavelengths).



Podemos distinguir alrededor de $N^3 = 10,000,000$ de colores, por lo que $N = \sqrt[3]{10,000,000} \approx 215.44$.

Por lo que si consideramos $N = 216 \Rightarrow N^3 = 10,077,696$

Así, si usamos una versión discretizada del cubo



con la $N = 216$, cada punto en el cubo se ve como

$$x = (r_i, g_j, b_k)$$

con $i, j, k = 1, \dots, 216$

Por lo que podemos denotar cada color distingible como un punto.

Así, al agregar una dimensión más, tendríamos cada color como un punto de la forma

$$x = (r_i, g_j, b_k, w_n)$$

Con $i, j, k, n = 1, \dots, 216$

$$\text{Teniendo apox. } N^3 \cdot N = (10,077,696)(216)$$

$$\approx 2.176782336 \times 10^9 \text{ colores distinguibles.}$$

(4)

4. [2 pts] An affine transform in the 2D plane is any linear transformation between *homogeneous coordinates* that has the following matrix form:

$$M = \begin{pmatrix} A & t \\ 0 & 1 \end{pmatrix},$$



with A a 2×2 invertible matrix and t a 2×1 vector. Demonstrate that **parallelism is conserved** under this kind of transformations, i.e. show that the image of two parallel lines through A are still parallel.

Consideramos dos rectas en su forma vectorial

$l_1 : (x, y) = p + k\vec{u}$; donde p es punto, k escala y \vec{u} vector

$l_2 : (x, y) = q + k\vec{u}$; q un punto

Estas dos rectas son paralelas, ahora, ambas rectas en coordenadas homogéneas, multiplicaremos por la matriz M

$$\begin{aligned} l_1' &= M(p + k\vec{u}) = Mp + M(k\vec{u}) = Mp + k(M\vec{u}) \\ &= \left(\begin{array}{cc} A & t \\ 0 & 1 \end{array} \right) \begin{pmatrix} p_1 \\ p_2 \\ 1 \end{pmatrix} + k \left(\begin{array}{cc} A & t \\ 0 & 1 \end{array} \right) \begin{pmatrix} u_1 \\ u_2 \\ 0 \end{pmatrix} \\ &= \underbrace{\left(\begin{array}{cc} A(p_1) & t \cdot 1 \\ 0 & 1 \end{array} \right)}_{Mp} + k \underbrace{\left(\begin{array}{cc} A(u_1) & 0 \\ 0 & 0 \end{array} \right)}_{M\vec{u}} \end{aligned}$$

Similármamente para la 2da recta

$$\begin{aligned} l_2' &= M(q + k\vec{u}) = Mq + M(k\vec{u}) = Mq + k(M\vec{u}) \\ &= \left(\begin{array}{cc} A & t \\ 0 & 1 \end{array} \right) \begin{pmatrix} q_1 \\ q_2 \\ 1 \end{pmatrix} + k \left(\begin{array}{cc} A & t \\ 0 & 1 \end{array} \right) \begin{pmatrix} u_1 \\ u_2 \\ 0 \end{pmatrix} \\ &= \underbrace{\left(\begin{array}{cc} A(q_1) & t \cdot 1 \\ 0 & 1 \end{array} \right)}_{Mq} + k \underbrace{\left(\begin{array}{cc} A(u_1) & 0 \\ 0 & 0 \end{array} \right)}_{M\vec{u}} \end{aligned}$$

Notamos que el vector dirección $M\vec{u}$ se conserva en la aplicación de las transformaciones en ambas rectas paralelas, por lo que el **parallelismo** se conserva



5. [2 pts] Let \mathbf{l} be the 3×1 homogeneous coordinates of a line in the Euclidean plane. This means that a point \mathbf{p} in the plane (expressed in homogeneous coordinates) belongs to this line iff

$$\mathbf{l}^T \mathbf{p} = 0.$$

Suppose that we apply some linear, invertible transform \mathbf{T} (represented as a 3×3 matrix) to the homogeneous coordinates \mathbf{p} (that transform them into $\mathbf{p}' = \mathbf{T}\mathbf{p}$). Demonstrate that the image points obtained this way from points on \mathbf{l} are *also belonging to a line \mathbf{l}'* , whose parameters are given by:

$$\mathbf{l}' = \mathbf{T}^{-T} \mathbf{l}.$$

Sea \mathbf{p} un punto en la recta \mathbf{l} , i.e., $\mathbf{l}^T \mathbf{p} = 0$, y sea \mathbf{T} una transformación invertible, $T_{3 \times 3}$, queremos probar que si al aplicar \mathbf{T}^a a \mathbf{p} , entonces $\mathbf{p}' = \mathbf{T}\mathbf{p}$ está en $\mathbf{l}' = \mathbf{T}^{-T} \mathbf{l}$.

Como $\mathbf{p} \in \mathbf{l} \Rightarrow \mathbf{l}^T \mathbf{p} = 0$

luego al aplicar \mathbf{T} a \mathbf{l} , tenemos $\mathbf{l}' = \mathbf{T}^{-T} \mathbf{l}$
y \mathbf{T}^a a \mathbf{p} , $\mathbf{p}' = \mathbf{T}\mathbf{p}$

$$\Rightarrow (\mathbf{l}')^T \mathbf{p}' = (\mathbf{T}^{-T} \mathbf{l})^T (\mathbf{T}\mathbf{p})$$

$$= (\mathbf{l}^T \mathbf{T}^{-1}) \mathbf{T}^T \mathbf{p}$$

$$= \mathbf{l}^T (\mathbf{T}^{-1} \mathbf{T}) \mathbf{p}$$

$$= \mathbf{l}^T \mathbf{p} = 0 \quad \text{como se requería.}$$

↑
es cero

por definición
pues \mathbf{p} está en \mathbf{l} .

Algunas notas que usé para entender el problema 5.

Primero, sea $T = \begin{pmatrix} A_{2 \times 2} & B \\ 0_{2 \times 2} & I \\ C & D \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ una matriz en bloques.

para calcular su inversa, tenemos que

$$T^{-1} = \begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} & -A^{-1}BD^{-1} \\ 0 & D^{-1} \end{pmatrix}$$

esta es $= \begin{pmatrix} A^{-1} & -A^{-1}C \cdot 1 \\ 0 & 1 \end{pmatrix}$ $\begin{pmatrix} 1 & 4 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$

Luego su transpuesta

$$T^{-\top} = \begin{pmatrix} A^{-\top}_{2 \times 2} & 0_1 \\ -C^TA^{-\top}_{2 \times 2} & 1_{1 \times 1} \end{pmatrix} \checkmark \leftarrow \text{la transpuesta inversa}$$

Por lo que queremos probar que

$$\lambda^{1\top} p = 0 \Leftrightarrow (T^{-\top}\lambda)^T(T_p) = \lambda^T \cdot T^{-1}(T_p) = \lambda^T p$$

Por lo que

$$T^{-\top}\lambda = \begin{pmatrix} A^{-\top}_{2 \times 2} & 0_1 \\ -C^TA^{-\top}_{2 \times 2} & 1_{1 \times 1} \end{pmatrix} \begin{pmatrix} q \\ b \\ 1 \end{pmatrix} = \begin{pmatrix} A^{-1}(\frac{q}{b}) & 0 \\ -C^TA^{-1}(\frac{q}{b}) & 1 \end{pmatrix}$$

Luego al trasponer

$$(T^{-\top}\lambda)^T = \begin{pmatrix} A^{-1}(\frac{q}{b})_{2 \times 2} & [-C^TA^{-1}(\frac{q}{b})]^T_{2 \times 1} \\ 0_{3 \times 2} & 1_{1 \times 1} \end{pmatrix}$$

Ahora por T_p

$$(T^{-1}\lambda)^T T_p$$

