Exercise *Computer Vision (CSCI 4220U)*

Please hand in this paper to the instructor before the end of the lecture.

Name: ____

Student number: ________________________ Date: ________________

Q. Consider a world point (X, Y, Z) that is imaged by a camera of focal length f at (x, y). Construct the pinhole camera equations that describe the relationship between (X, Y, Z) and (x, y).

Q. Set up an example camera system and write down its intrinsic and extrinsic ~~cameras~~ matrices.

Q. Show that lines are mapped to lines when passed through an *affine* transformation seen below:

$$\mathbf{x}_{\text{new}} = \mathbf{A}\mathbf{x}_{\text{old}} + \mathbf{t}.$$

Q. Construct a line between two points $(4, 1)$ and $(4, 2)$.

Q. Consider an 512×512 grayscale image where each pixel is stored using 8-bits. How much memory is needed to store the Guassian Pyramid of this image.

Q. Given an image $\mathbf{I} \in \mathbb{R}^{128 \times 128}$, provide a recipe for constructing its Laplacian pyramid.

Q. Compute the gradient of the following function at $(1, 2, 3)$:

$$f(x, y, z) = x^3 y + yz + 4z - 2.$$

?

Q. Give a one line definition of a periodic function.

Q. Use Taylor Series expansion, brightness constancy and small motions to write down the equation for solving optical flow.

Q. Write down the loss that is commonly used for linear regression problems?

Q. Write down the loss that is commonly used for logistic regression problems?

Q. Consider the following model:

$$f(x_1, x_2; \theta_1, \theta_2, \theta_3) = \theta_1 x_1 + \theta_1 \theta_2 x_2 + \theta_3.$$

not a linear regression model

How would you fit this model to data using linear regression?

Q. Briefly list the difference between MLP and CNNs in terms of their ability to deal with image data.

Q. Define the property of *spatial coherence* within the context of computer vision.

Q. Why is it more advantageous to blur an image using a Gaussian filter than, say using an averaging filter.

Q. How does the choice of similarity metric (e.g., cross-correlation, sum of squared differences, or normalized cross-correlation) impact the accuracy and robustness of template matching in image processing?

Q. Say you are asked to design a system to classify images containing cats. Cats are clearly visible in some images. Other images have a lot of clutter and sometimes cats are partially occluded. Then there are images that do not contain cats. You have two options to design a system: 1) a system that uses a global image features and 2) a system that uses local image features. Which of the two systems will you select, and why?

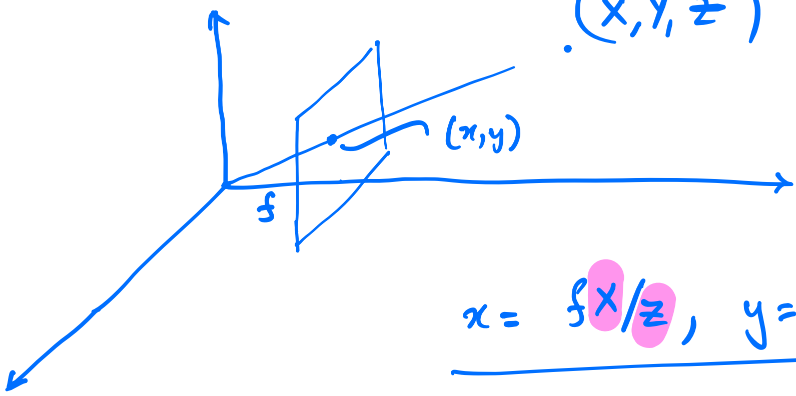
Q. What are the main parameters of RANSAC?

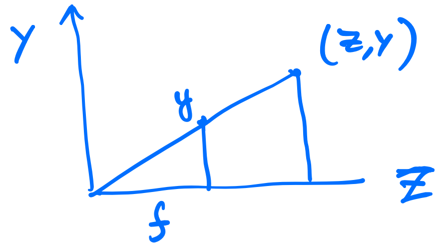
iterations

inlier threshold

model parameters.

Q.1





$$\underline{x = fX/Z, \quad y = fY/Z} \leftarrow ①$$

Q.2 Focal length = $f \in \mathbb{R}$

Camera Centre = $\vec{C} \in \mathbb{R}^3$

Rotation matrix = $R \in \mathbb{R}^{3\times3}$

Intrinsic matrix = $\underbrace{\begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix}} \in \mathbb{R}^{3\times3}$

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

$\downarrow$ Cartesian

$$\begin{bmatrix} fX/Z \\ fY/Z \end{bmatrix} \rightarrow ①$$

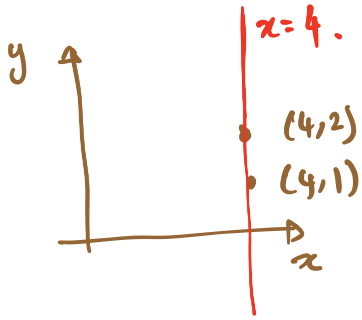
Extrinsic matrix (change of $\frac{\text{basis}}{\text{coordinate}}$ systems)

$$[\, R \mid -RC \,] \in \mathbb{R}^{3\times4}$$

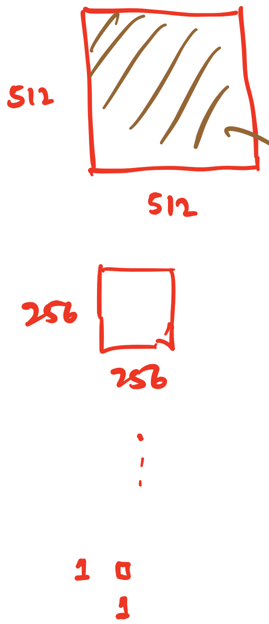
$$= \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

Q.3 skip

Q.4



Q.5



$(512 \times 512) \times 8$

$\left(2^9\right)^2 \longrightarrow \left[\sum_{i=0}^{9} \left(2^i\right)^2\right] \times 8$

Q.6

$G_0 = I \in \mathbb{R}^{128\times128}$

$G_1 \in \mathbb{R}^{64\times64}$ $\xrightarrow{\text{upsample}}$ $\hat{G}_0$:

$G_2 \in \mathbb{R}^{32\times32}$

$G_3 \in \mathbb{R}^{16\times16}$

$G_4 \in \mathbb{R}^{8\times8}$

$G_5 \in \mathbb{R}^{4\times4}$

$G_6 \in \mathbb{R}^{2\times2}$

$G_7 \in \mathbb{R}$

$L_0 = G_0 - \hat{G}_0$

L_1

L_2

L_3

L_4

L_5

L_6

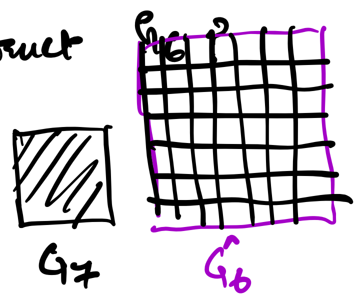
G_7

$\hat{G}_i = \text{upsample}(G_{i+1})$

$L_i = G_i - \hat{G}_i$

Q. Given L_6 and G_7 construct G_6?

$\hat{G}_6 = \text{upsample}(G_7)$



$G_6 = \hat{G}_6 + L_6$

L_i

G_{i+1}

$i = 3$

L_3

G_4

Q.

Q. $f(x) = f(x + \varepsilon)$ $\quad \varepsilon > \emptyset$: $\quad \varepsilon \neq 0$

Q. $I(x, y, t) = \underbrace{I(x + u, y + v, t+1)}$

$$I(x+u, y+v, t+1) = I(x,y,t) + u I_x + v I_y + I_t$$

$\Rightarrow$ $u I_x + v I_y = - I_t$

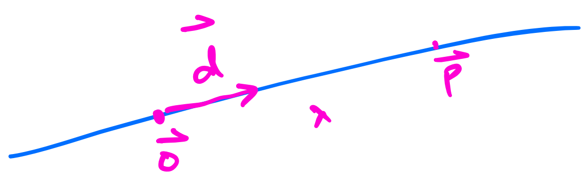
$$\frac{\partial \; I(x,y,t)}{\partial x}$$

$\vec{x}_{new} = A \vec{x}_{old} + \vec{t}.$ $\quad \vec{x}_{new}, \vec{x}_{old}, \vec{t} \in \mathbb{R}^3 \quad A \in \mathbb{R}^{3\times 3}$

Eq. of a line in parametric form:

$\vec{p} = \vec{o} + \vec{d}\lambda$ $\quad \vec{p} \in \mathbb{R}^3 \quad \vec{o}, \vec{d} \in \mathbb{R}^3$, $\lambda \in \mathbb{R}$

$\lambda \in [-\infty, \infty]$



new "curve"

line

$\vec{x} = A \vec{p} + \vec{t}$

$$\vec{x} = A(\vec{o} + \vec{d}\lambda) + \vec{t}$$

$$= A\vec{o} + A\vec{d}\lambda + \vec{t}$$

$$= \vec{o}' + \vec{d}'\lambda + \vec{t}$$

$$= (\vec{o}' + \vec{t}) + \vec{d}'\lambda$$

$$\vec{x} = \underbrace{\vec{o}'' + \vec{d}'\lambda}_{\text{line}}$$

$\vec{o}' \in \mathbb{R}^3$

$\vec{d}' \in \mathbb{R}^3$

$\vec{o}'' \in \mathbb{R}^3$