CS512 - Assignment 1

Mohammadreza Asherloo
Department of Mechanical, Materials and Aerospace Engineering
Illinois Institute of Technology

October 10, 2020

Question 1.a:

$$\begin{bmatrix} U \\ V \end{bmatrix} = \frac{1}{z} \begin{bmatrix} -F & 0 \\ 0 & -F \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$
$$\begin{bmatrix} U \\ V \end{bmatrix} = \frac{1}{1} \begin{bmatrix} -10 & 0 \\ 0 & -10 \end{bmatrix} \begin{bmatrix} 3 \\ 2 \end{bmatrix} = \begin{bmatrix} -30 \\ -20 \end{bmatrix}$$

Question 1.b:

Regarding the equations, both models are the same but the model where image plane is front of COP, doesn't need the inversion. The model with image plane behind the COP corresponds better to physical model and the other model is good for helping to visualize and making pictures.

Question 1.c:

When the focal length is getting bigger, the projection of object gets bigger, too.
On the other hand, if the Z gets bigger, the projection of object gets smaller.

Question 1.d:

2D: $\big(1, 1\big)$ === 2DH: $\big(1, 1, 1\big)$ === 2DH: $\big(2, 2, 2\big)$

Question 1.e:

2DH: $\big(1, 1, 2\big)$ === 2D: $\big(0.5, 0.5\big)$

Question 1.f:

It means that this point is in the infinity direction.

Question 1.g:

Using homogeneous coordinates makes it possible to use linear equations for projection. For example, using 2DH coordinates instead of 2D coordinates.

Question 1.h:

$M = 3 \times 4$
$K = 3 \times 3$
$I = 3 \times 3$
$0 = 3 \times 1$

Question 1.i:

$$\begin{bmatrix} U \\ V \\ W \end{bmatrix} = \begin{bmatrix} F & 0 & 0 & 0 \\ 0 & F & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} U \\ V \\ W \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 1 & 2 & 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 18 \\ 46 \\ 10 \end{bmatrix} \text{(2DH)} === \begin{bmatrix} 1.8 \\ 4.6 \end{bmatrix} \text{(2D)}$$

Question 2.a:

$$\begin{bmatrix} U \\ V \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & Tx \\ 0 & 1 & Ty \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} U \\ V \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \\ 1 \end{bmatrix} === \begin{bmatrix} 3 \\ 4 \end{bmatrix} \text{(After translation)}$$

Question 2.b:

$$\begin{bmatrix} U \\ V \\ 1 \end{bmatrix} = \begin{bmatrix} Sx & 0 & 0 \\ 0 & Sy & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

$$\begin{bmatrix} U \\ V \\ 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} === \begin{bmatrix} 2 \\ 2 \end{bmatrix} \text{ (After translation)}$$

Question 2.c:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} Cos\theta & -Sin\theta \\ Sin\theta & Cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} Cos(45) & -Sin(45) \\ Sin(45) & Cos(45) \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1.4 & -1.4 \\ 1.4 & 1.4 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 2.8 \end{bmatrix}$$

Question 2.d:

R(p,u) = T(p)Ru(θ)T(-p)

$$\text{R(p,u)} = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} Cos(45) & -Sin(45) & 0 \\ Sin(45) & Cos(45) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\text{P' = R(p,u) P} = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} Cos(45) & -Sin(45) & 0 \\ Sin(45) & Cos(45) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} Cos(45) & -Sin(45) & 0 \\ Sin(45) & Cos(45) & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 2 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ -2.8 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ -0.8 \\ 1 \end{bmatrix}$$

$$=== \begin{bmatrix} 0 \\ -0.8 \end{bmatrix} \text{ (After rotation)}$$

Question 2.e:

P' = TRP

Question 2.f:

$$\text{M} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} S & 0 \\ 0 & 1 \end{bmatrix} \text{ (This is a scale transformation matrix by (3,2))}$$

Question 2.g:

M = $\begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} I & T \\ 0 & 1 \end{bmatrix}$ (This is a translate transformation matrix by (1, 2))

Question 2.h:

$M^{-1} = \begin{bmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix}$

Question 2.i:

$M^{-1} =$ T(-1,-2)R(-45) $= \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} Cos(-45) & -Sin(-45) & 0 \\ Sin(-45) & Cos(-45) & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1.4 & 1.4 & -1 \\ -1.4 & 1.4 & -2 \\ 0 & 0 & 1 \end{bmatrix}$

Question 2.j:

$V_1 \, . \, V_2 = 0$, $V_1 = \begin{bmatrix} 1 \\ 3 \end{bmatrix}$, $V_2 = \begin{bmatrix} i \\ j \end{bmatrix}$

i + 3j = 0 $\Rightarrow$ i = -3j $\Rightarrow$ $V_2 = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$

Question 2.k:

projection $= \frac{V_1.V_2}{V_2.V_2} V_2 = \frac{17}{29} V_2 = \begin{bmatrix} \dfrac{34}{29} & \dfrac{85}{29} \end{bmatrix}$

Question 3.a:

We need this general model to eliminate the problem of using transformations in the same coordinate system because the camera can move.

Question 3.b:

$M_{C \longleftarrow W} = \begin{bmatrix} R^T & -R^T T \\ 0 & 1 \end{bmatrix}$

Question 3.c:

$$R^T = \begin{bmatrix} \hat{x}^T \\ \hat{y}^T \\ \hat{z}^T \end{bmatrix}$$

Question 3.d:

R^*: This is the rotation matrix of world with respect to camera
T^*: This is the translation matrix of world with respect to camera

Question 3.e:

$$M_{i \longleftarrow C} = \begin{bmatrix} k_u & 0 & 512 \\ 0 & k_v & 512 \\ 0 & 0 & 1 \end{bmatrix}$$

Question 3.f:

K^*: This part contains the intrinsic parameters of the camera (focal length)
R^* and T^*: These parts contain the extrinsic parameters of the camera (rotation and translation)

Question 3.g:

When we want accuracy in the scale of millimeter we need to use the skew parameter.

Question 3.h:

The radial lens distortion is modeled which includes a linear and a quadratic part and depends on the distance from center of the image. After this modeling of radial lens distortion we can correct it on the image by adding the matrix below to transformation matrix:
$$\begin{bmatrix} \frac{1}{\lambda} & 0 & 0 \\ 0 & \frac{1}{\lambda} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$
The complication of adding this parameter to transformation matrix is that this parameter changes for every pixel and is not constant.

Question 3.i:

Weak perspective camera: in these cameras we can't see any perspective of the scene in the image.
Affine camera: in these cameras the transformation matrix consists of 8 arbitrary numbers and the last row will be (0 0 0 1) and the whole matrix is a computational model.

Question 4.a:

Surface radiance: this parameter shows that how much light is being reflected from a particular surface
Image irradiance: this parameter shows that how much light is being received by the image.

Question 4.b:

Image irradiance = surface radiance $\times \frac{\pi}{4} \times \left(\frac{d}{f}\right)^2 \times (Cos\alpha)^4$
Where d is diameter of lens, f is focal length and α is the angle between principal axis and surface normal.

Question 4.c:

Albedo of a surface means the reflection coefficient which is between 0 and 1 and higher the albedo, higher the amount of light being reflected from that surface.

Question 4.d:

RGB color model is used because human vision has receptors of Red, Green and Blue lights so it makes it easier to map colors in computer vision.

Question 4.e:

On this line going from origin to (1 1 1) we will have different shades of gray color going from dart gray to light gray.

Question 4.f:

We have a target color and we change the R,G and B knobs to match the color to the target color. after this process we can have different wavelength of Red, Green and Blue light for each color and after plotting these points we can have a mapping of colors by RGB model.

Question 4.g:

Y is the luminance and is useful for black and white broadcasting. By receiving luminance the broadcaster can show different shades of gray good enough for black and white contents.

Question 4.h:

Advantage of LAB model is that it takes euclidean distance of two colors and measures the distance of them to a reference color and tells us which color is closer to the reference. This process corresponds to the human perception of colors.

Question 5.a:

For computing the SNR we need to compute the variance of the signal and variance of the noise. For variance of the noise we have two options:
1. Compute the variance of image on different images of the same scene.
2. Compute the variance of a region of the image which is uniform and extrapolate it to the whole image.

Question 5.b:

Gaussian noise is a kind of noise in the form of normal distribution but impulsive noise has random independent values.
Median filter is better for impulsive noise because it will ignore the outliers.

Question 5.d:

Question 5.e:

First approach is zero padding in which we add desired rows and columns of zero to the image.
Second approach is mirroring in which we add the same row and column of the edge to the image.
Third approach is ignoring the values of padding.

Question 5.f:

Sum of all the elements should be 1. The reason is that we want to give the same weight to all of the image pixels and take the average of the matrix.

Question 5.g:

We can convolve the image with a 1D Gaussian filter along the rows and then after that, do it along the columns.
No, we can't do this process for all kind of filters because of the nature of Gaussian filters which are symmetrical along the rows and columns.
Using 1D filter twice is more efficient.

Question 5.h:

The size of the filter should be 13 by 13.

Question 5.i:

After each layer, we do some pullings to cut the size of the image to quarter of the last layer.
We use these pyramids to implement multiscale analysis of the image and overcome the aliasing problem of large images.
The added amount of calculations is about one third of the original calculations.

Question 5.j:

In the Laplacian pyramid, after down sampling the image, we will up sample it again to the original size before down sampling and convolve it with the filter to see what we lost during the down sampling.

Question 6.a:

Edge detection is important because edges can give us precise information about the contents of the image.
It has to correspond to the scene elements, it should be invariant with illumination, pose or other properties of elements and it should be reliable.

Question 6.b:

First we smooth the image to prevent the derivatives to detect noises.
Second we enhance the edges to amplify them for detection by derivatives.
Third we detect them.
Fourth we localize them to find their location in the image.

Question 6.c:

X derivative: $\begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}$

Y derivative: $\begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}$

Gradient in the images shows the amount of change of intensity in every pixel.
We cab use image gradients to detect the change of intensity and as a result, detect the edges.

Question 6.d:

Sobel filter is the result of convolution of a smoothing filter and a derivative filter:
Smoothing filter * X derivative filter= $\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} * \begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$

Question 6.e:

For the X derivative we can calculate the convolution of image * X Gaussian derivative * Y Gaussian smoothing.
For the Y derivative we can calculate the convolution of image * Y Gaussian derivative * X Gaussian smoothing.
We can compute the Gaussian and Gaussian derivative filters with these equations:
$G(x) = e^{-\frac{x^2}{2\sigma^2}}$
$G'(x) = -\frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}}$

Question 6.f:

We can use first order derivative to detect the edge by a threshold and we can have more accurate information about the location of the edge.

By using second order derivative we can have the exact location of the edge by zero crossing technique meaning that wherever the second derivative is zero, we have an edge.

Question 6.g:

$$\begin{bmatrix} 0.001 & 0.016 & 0.053 & 0.07 & 0.053 & 0.016 & 0.001 \\ 0.016 & 0.1 & 0.22 & 0.26 & 0.22 & 0.1 & 0.016 \\ 0.053 & 0.22 & 0 & -0.6 & 0 & 0.22 & 0.053 \\ 0.07 & 0.26 & -0.6 & -2 & -0.6 & 0.26 & 0.07 \\ 0.053 & 0.22 & 0 & -0.6 & 0 & 0.22 & 0.053 \\ 0.016 & 0.1 & 0.22 & 0.26 & 0.22 & 0.1 & 0.016 \\ 0.001 & 0.016 & 0.053 & 0.07 & 0.053 & 0.016 & 0.001 \end{bmatrix}$$

This filter is the result of convolution of a Laplacian and Gaussian kernels so it does some smoothing (Gaussian) to reduce noise and uses second derivative (Laplacian) to detect the edges by zero crossing technique.

Question 6.h:

Canny edge detector uses non-max suppression to pick the best pixel for edge and uses hysteresis threshold detect the possible edge neighbors, thus, the results of canny method are much more reliable.

Question 6.i:

Non-suppression: with this technique we can find the maximum gradient magnitude in the gradient direction to pick the best pixel for edge detection in the edge region.
Hysteresis threshold: with this technique after finding the edge, we can look in the neighborhood to find the possible edges and track the edge