ME5405 Machine Vision

Sample Solution for Exercise 2: Image Acquisition

There may be more than one correct answer to the question. Email me at mpecck@nus.edu.sg if you have found any errors in the solution. – CK Chui

Question 1: Suppose that the accuracy of the system is defined by the distance travelled in the period of 1 frame. Given that the object (postal parcel) on the conveyer belt is moving at a speed of 0.3 m/s. What is the accuracy of this system?

Answer
Given that the speed is 0.3 m/s and assuming that the camera is acquiring images at a frame rate of 30 frames per second.
Accuracy = 0.3 m/s x 1/30 = 0.01 m = 1 cm

[Note that frame rate of the camera is not given. 30 frames per second is a good assumption for industrial application.]

Question 2: Assume that the camera can move horizontally. Write an algorithm to "center" the parcel in the image.

Answer

(There can be more than one answer.)

Step 1: Determine the edge of the parcel.
Step 2: If one edge is not entirely in the acquired image, move the camera horizontally towards the missing part of the edge.
Step 3: Repeat Step 2 for all edges
Step 4: When entire object is visible in the image, compute the object center, and align the image accordingly.

Question 3: Write an algorithm to determine the area of the parcel in the image (in terms of pixels).

Answer
Assumption: the acquired image is a gray scale image.

Step 1: Convert the gray scale image to a binary image using brightness thresholding.
Step 2: The black pixels in the resultant binary image represent the object (parcel). Calculate the number of black pixels.

Question 4: The depth of field determines whether the text on the parcel can be read clearly (or sharply). Explain how you can determine quantitatively whether the text on the image will be sharp or blur.

<u>Answer</u>

Assumption: there is sufficient illumination during image acquisition to capture the text on the parcel.

The depth of field D can be determined from the pixel size, f-stop and magnification. As long as the object plane is within the depth of field, the object will appear sharp.

Suppose that h is the height of the camera above the top of the parcel. Since h is fixed, and pixel size, f-stop and magnification are given in the camera setting, we can quantitatively determine D and hence, whether the text on the image will appear sharp.

Question 5: Given that the camera comprises of a rectangular 8.8 x 6.6 mm CCD sensor array with 768 x 493 pixels and a lens with focal length = 300 mm. The frame rate is 30 frames per second.

If the magnification is 10, what is the size of one pixel on the object plane? At what distance, the object will appear completely sharp on the image? Given that the focus length is 300 mm.

<u>Answer</u>

Magnification = 10

Pixel size = CCD array size / CCD no. of pixels
= 8.8 mm / 768 by 6.6 mm / 493
= 11.5 by 13.4 μm

Pixel size on the object = pixel size / magnification
= 11.5 by 13.4 μm / 10
= 1.15 by 1.34 μm

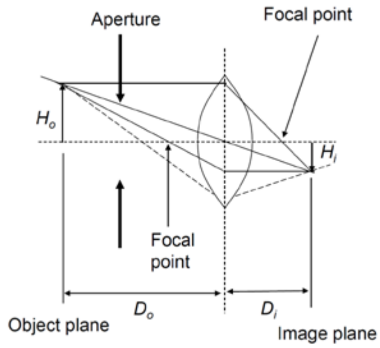
The object will appear completely sharp at the sharp object plane, i.e.

(focal length) x (1 + 1/magnification) = 300 mm x 1.1

Question 6: A CCD camera sensor of dimension 14 x 14 mm, and having 2048 x 2048 elements, is focused on a square, flat area, located 0.5 m away. Suppose that the camera is equipped with a 35-mm lens. How many elements per mm will this camera be able to resolve?

Answer

From the geometry of the figure given in the note:



D_o = 500 mm; D_i = 35 mm; H_i = 14/2 = 7 mm.
H_o/(500 mm) = (7 mm)/(35 mm)
H_o = 100 mm; the target size is 100 mm on the side.
Since there is a total of 1024 elements per line, the resolution of 1 line is 1024/100 == 10 elements/mm.
Number of elements or lines per mm = 10

Question 7: Where will the image appear if the object has been re-positioned to 2.5 cm from the lens?

Describe the image.

Answer

When the object distance d_o is 2.5 cm,

$$\frac{1}{2.5}+\frac{1}{d_i}=\frac{1}{7.5} \quad \rightarrow d_i=-3.75\text{cm}$$

$$m=-\frac{-3.75}{2.5}=+1.5$$

Hence, the image is virtual, erect, magnified and located 3.75 cm in front of the lens.

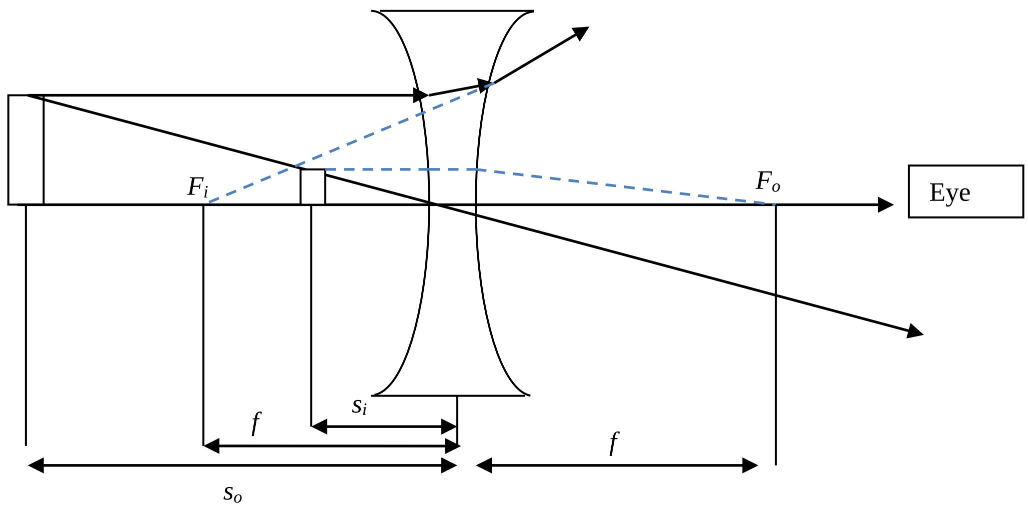
Question 8: What is the type of lens? Suppose that the focus length of the lens is *f*. Express the object and image distances in terms of *f*. Construct a ray diagram.

<u>Answer</u>

Since the image is erect but minified, the lens must be a diverging lens.

$$\mathrm{m} = \frac{1}{3} = -\frac{s_i}{s_o}$$

$$\frac{1}{(-3s_i)} + \frac{1}{s_i} = \frac{1}{f} \rightarrow s_i = \frac{2f}{3}$$



<u>What is the height of retinal image of the object?</u>

Suppose that the normal height of the object is X cm. Note that the distance between focal center of the lens and the retina varies from approximately 17 mm to about 14 mm. Assuming that you are looking at the image X/3 cm tall at d cm away, and the distance between focal center of the lens and the retina is 15 mm.
If x is the height of the retinal image in mm,
X/3d = x/15, i.e. x = 5X/d.

Question 9: Given an optical system comprises of a two-lens combination (compound lens) of focal lengths f_1 and f_2 respectively. The two lenses are separated by a distance d. Suppose that T_1 and T_2 are the translation matrices for the ray in air before and after the two-lens system. T_d is the translation matrix for a ray travelling between the two lenses. S_1 and S_2 are the lens matrix respectively.

Determine the system matrix **M**.

<u>Answer</u>

$$\begin{aligned}
\begin{pmatrix} r_i \\ v_i \end{pmatrix} &= \mathbf{T}_2\mathbf{S}_2\mathbf{T}_d\mathbf{S}_1\mathbf{T}_1\begin{pmatrix} r_0 \\ v_0 \end{pmatrix} \\
&= \begin{pmatrix} 1 & d_2 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -1/f_2 & 1 \end{pmatrix}\begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -1/f_1 & 1 \end{pmatrix}\begin{pmatrix} 1 & d_1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} r_0 \\ v_0 \end{pmatrix}
\end{aligned}$$

$$\begin{aligned}
\mathbf{M} &= \begin{pmatrix} 1 & d_2 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -1/f_2 & 1 \end{pmatrix}\begin{pmatrix} 1 & d \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -1/f_1 & 1 \end{pmatrix}\begin{pmatrix} 1 & d_1 \\ 0 & 1 \end{pmatrix} \\
&= \ldots
\end{aligned}$$

Question 10: Image subtraction is used often in industrial applications for detecting missing components in product assembly. The approach is to store a "golden" image that corresponds to a correct assembly; this image is then subtracted from incoming images of the same product. Ideally, the differences would be zero if the new products are assembled correctly. Difference images for products with missing components would be nonzero in the area where they differ from the golden image.

Suppose that $g(x, y)$ denotes the golden image, and $f(x, y)$ denotes any input image acquired during routine operation of a machine vision system. The resulting image based on simple difference is $d(x, y) = g(x, y) - f(x, y)$.

$f(x, y)$ is "close enough" to the golden image if all the pixels in $d(x, y)$ fall within a specified threshold band $[T_{min}, T_{max}]$ where T_{min} is negative and T_{max} is positive. Usually, the same value of threshold is used for both negative and positive differences, so that we have a band $[-T, T]$ in which all pixels of $d(x, y)$ must fall in order for $f(x, y)$ to be declared acceptable.

What conditions do you think have to be met in practice for this method to work?

Answer

In practice, there are three fundamental factors that need tight control for difference based inspection to work:

(1) Proper registration,

(2) Controlled illumination, and

(3) Noise levels that are low enough so that difference values are not affected appreciably by variations due to noise.

The first condition basically addresses the requirement that comparisons be made between corresponding pixels. Two images can be identical, but if they are displaced with respect to each other, comparing the differences between them makes no sense. Often, special markings are manufactured into the product for mechanical or image-based alignment.

Controlled illumination or lighting obviously is important because changes in illumination can affect dramatically the values in a difference image. One approach used often in conjunction with illumination control is intensity scaling based on actual conditions. For example, the products could have one or more small patches of a tightly controlled color, and the intensity (and perhaps even color) of each pixels in the entire image would be modified based on the actual versus expected intensity and/or color of the patches in the image being processed.

Finally, the noise content of a difference image needs to be low enough so that it does not materially affect comparisons between the golden and input images. Good signal strength goes a long way toward reducing the effects of noise. Another and sometimes complementary approach is to implement image processing techniques such as image averaging to reduce noise.