**Assignment 1 Solution  
Computer Vision (CS-559)**

**Name: Dhaval Harish Sharma  
Red ID: 824654344**

1. **(8 points) Consider a square region of size m by m pixels in each of two images f1(x,y) and f2(x,y). The RGB colors of these two regions are denoted by R1(x,y), G1(x,y), B1(x,y), and R2(x,y), G2(x,y), B2(x,y), where (x0,y0) is the position of the upper left corner of the square region, and (x,y) is the position of a typical pixel in the region, i.e. x=x0, x0+1, …, x0+m-1and y=y0, y0+1, …, y0+m-1 . Propose a measure for the similarity (not identically) between the colors of these two regions. Express the measure in analytical form and explain why your measure is appropriate. If two regions are very similar by your measure, will that mean that they are visually similar? Demonstrate your findings with the aid of two regions of an image.**

**Solution:**The measure for finding the similarity between the colours of two images is based on the average of the differenced value of pixels of the regions in both the images and then comparing that average.

The basic idea behind the measure is the difference between corresponding pixel values which can be used to compare the similarity of the pixel values in the region. The output of this algorithm results in a particular value between 0 and 1. All of these results can be cumulated and a brief idea about the whole picture can be obtained.

The algorithm for the same is as follows:   
1. Initialize a variable called “sum” to value 0.  
2. Start with the first pixel from the first row and first column.  
3. Modulo the subtraction of pixel values from both the images and add it to variable “sum”.   
4. Then, increment the pixels in the next column and go to step 3. If there are no more columns, then increment the row by 1 and go to step 3.  
5. After getting the sum of all the corresponding pixels divide the result by the total number of pixels in the window of both the images and store it in the variable “average”.  
6. Get the final output by normalizing the value of the average that is; divide the average by number of gray values available in the image.  
7. In the same way, continue doing this process until all the RGB values of the pixels in the image are traversed.

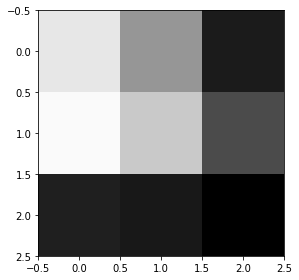
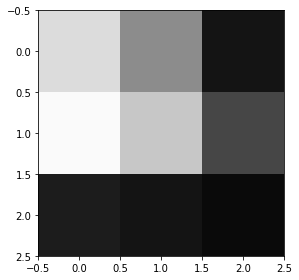
The output of the above algorithm is the measure of similarity between the two regions of the images. The range of the output is between 0 and 1. The closer the output is to 0, the more similar the images are. A value of 0 in the output means that the images are identical and a value of 1 in the output means the pixel values in both the images are inverse of each other.

If two regions are found similar by this algorithm, then it can be said that the regions are visually similar because both the regions are having almost similar kinds of pixel values.

Let us consider an example to understand this algorithm,  
We have two images which are almost similar in pixel values. If the pixel values consist of 3X3 pixels, we have two matrices as follows:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Image 1** | | | **Image 2** | | |
| 231 | 150 | 27 | 220 | 140 | 20 |
| 250 | 201 | 75 | 250 | 199 | 70 |
| 31 | 24 | 0 | 28 | 20 | 10 |

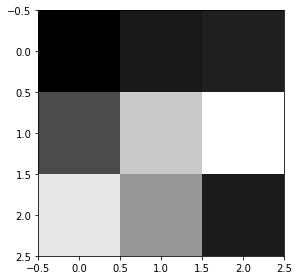
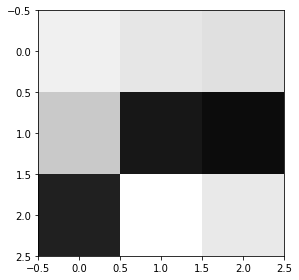
Sum = |231 – 220| + |150 – 140| + |27 – 20| + | 250 – 250| + |201 – 199| + |75 – 70| + |31 – 28| + |24 – 20| + |0 – 10| = 52  
Average = 52 / 9 = 5.7777  
Normalizing:   
**Measure Value = 5.7777 / 255 = 0.0226**

   
 Image 1 Image 2

Suppose we have two regions that are quite dissimilar than each other,

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Image 1** | | | **Image 2** | | |
| 0 | 24 | 31 | 240 | 230 | 224 |
| 75 | 201 | 255 | 201 | 23 | 12 |
| 231 | 150 | 27 | 33 | 255 | 233 |

Sum = 1695  
Average = 1695 / 9 = 188.3333  
Normalizing:   
**Measure Value = 188.3333 / 255 = 0.7385**

   
 Image 1 Image 2

1. **(20 points) In an automated manufacturing, inspection of circuit boards is to be done using a CCD camera. The individual imaging elements (photosites) each has a dimension of 5 by 5 m (micron) and the spacing between the elements is 1 m. The circuit boards are 60 by 60 mm, and defects appear as dark circular blobs with diameter of 0.4 mm or larger. The smallest defect must appear in the image as an area of at least 6 by 6 pixels. Assume that available lenses come with focal lengths of multiple of 25 mm, 35 mm and 50 mm, and the available camera resolutions are multiple of 256 by 256 pixels up to 2048 by 2048 pixels (4 Mpix). Manufacturing requirements dictate that distance between camera and the circuit board must be between 200 mm to 500 mm. The image of the board must occupy the whole image plane. You are to select the lens focal length and the minimum camera resolution (number of pixels) required. Show in reasonable details the analysis that lead to your answers.**

**Solution:**The first step is to calculate the number of pixels in the image plane. This can be found by mapping the size of error in circuit board to the number of pixels in the image plane. It is given in the question that an error of 0.4 mm in diameter appears as 6 X 6 pixels in the image plane. So, we calculate the total number of errors that can appear in a single row in the circuit board.

0.4 mm -> 6 pixels  
60mm -> (?)

Total number of pixels in the image plane = (60 \* 6) / 0.4 = 900 pixels

This gives us the size of the image plane which is 900 X 900 pixels, but we are given a set of resolutions of the image plane. So, the nearest resolution from the set of given resolutions is **1024 X 1024 pixels**.

Then, the next step is to find the actual area of the total image plane. We are given that an individual pixel’s size is 5 X 5 micrometer (0.005 X 0.005 mm) and the distance between individual pixels is 1 micrometer (0.001 mm), so we can calculate the distance of the image plane along x-axis which is as follows:

x = (1024 \* 0.005) + (1023 \* 0.001) = 6.143 mm  
In the same way,  
y = 6.143 mm

Now, we are also given that the image of the board should occupy the whole image plane. We can calculate the magnification factor  as follows:

 = Image size / Object size  
 = 6.143 mm / 60 mm  
 = 0.1024

According to the formula for focal length, we get

  
 = (0.1024 \* Z) / (0.1024 + 1)  
 f = 0.0928 Z  
Z = 10.7656 f

But, we are given that f should be multiple of 25, 35 and 50 whereas the value of Z should be between 200 and 500. This gives the possible values of f as follows:

[200 < Z < 500]

|  |  |
| --- | --- |
| **f** | **Z** |
| 25 | 269.14 |
| 35 | 376.796 |
| 50 | 538.28 |
| 70 | 753.592 |
| 75 | 807.42 |

Thus, there are only two values of f satisfy the value of Z which are **25 and 35**.

**In conclusion, we can say that the resolution of the image plane would be 1024 X 1024 and the focal length could be 25 or 35.**