Computer vision (CSC I6716)

Assignment 1

Prof. Zhigang Zhu
Sonali Shintre
ID: 7602
10/6/19

1. **How does an image change (e.g., objects' sizes in the image, field of view, etc.) if the focal length of a pinhole camera is varied?**

Focal length is the distance from the center of lens to the focal point. The image size changes linearly with the focal length f. As the focal length of the pinhole camera increases, the size of the image also increases and vice versa. The perspective camera x = f *X/Z, and y = f*Y/Z, has two effects: first, size and shape of object images changes linearly with focal length f and second, field of view (FOV) is increasing with decreasing f and vice versa.

2. **Give an intuitive explanation of the reason why a pinhole camera has an infinite depth of field**

If the camera's hole is big then large number of light rays that come into the camera. When they get in, they are not very organized, and that causes focus blur. So only one ray is passed through the hole onto the image plane. As the depth of the field increase its wideness decrease, since pinhole has a very small camera hole so the camera depth will be large and has infinite depth of field.

3. **In the thin lens model, 1/o + 1/i = 1/f, there are three variables, the focal length f, the object distance o and the image distance i (please refer to Slide # 19 of the Image Formation lecture). If we define Z = o-f, and z = i-f, please write two a few words to describe the physical meanings of Z and z, and then prove that Z*z = f*f given 1/o + 1/i = 1/f.**

f=focal length
o=object distance
i=image distance
we know that Z=o-f, difference between the object distance and the focal length
z=i-f difference between image distance and focal length

As the focal length increase from the lens, the length od the lens also increase and vise verse

-> i/o +1/i= i/f
-> (i+o) / (i* o)= i/f
-> (f * i) +(f * o) =(i*o)
-> (f * i) = (i*o) – (f*o)

-> (f * i) =o*(i-f)
-> (f * i)/o= z
Same with (f*o)/i =Z
So, Z *z= ((f * i)/o) * ((f*o)/i)
=> Z*z =f*f

4. **Prove that, in the pinhole camera model, three collinear points (i.e., they lie on a line) in 3D space are imaged into three collinear points on the image plane. You may either use geometric reasoning (with line drawings) or algebra deduction (using equations)**

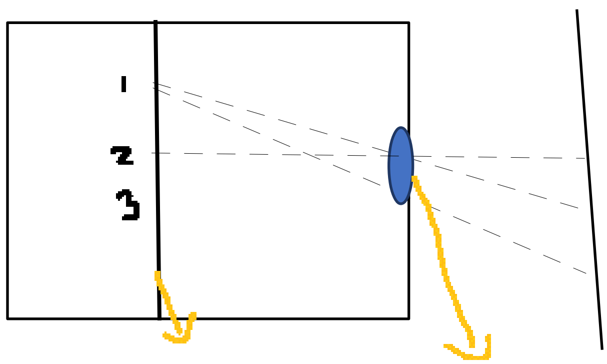
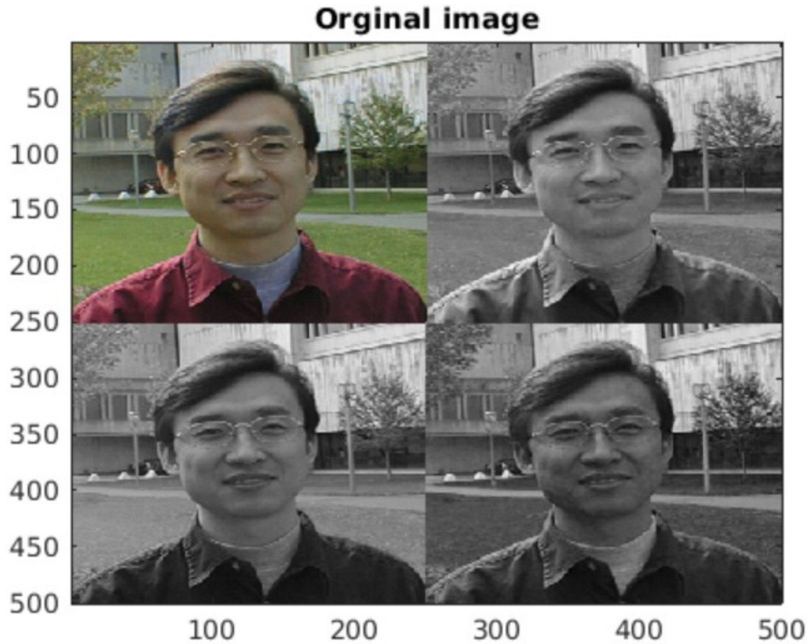


Image Plane **pinhole**

All the lights coming through the pinhole are depicts after the pin hole on the image plane, so the lines are straight before and after the pinhole.
3 collinear points in coordinates 1,2, 3 form a matrix whose determinant is 0. We can apply the perspective projection to each point to get the three points x1, x2, x3.

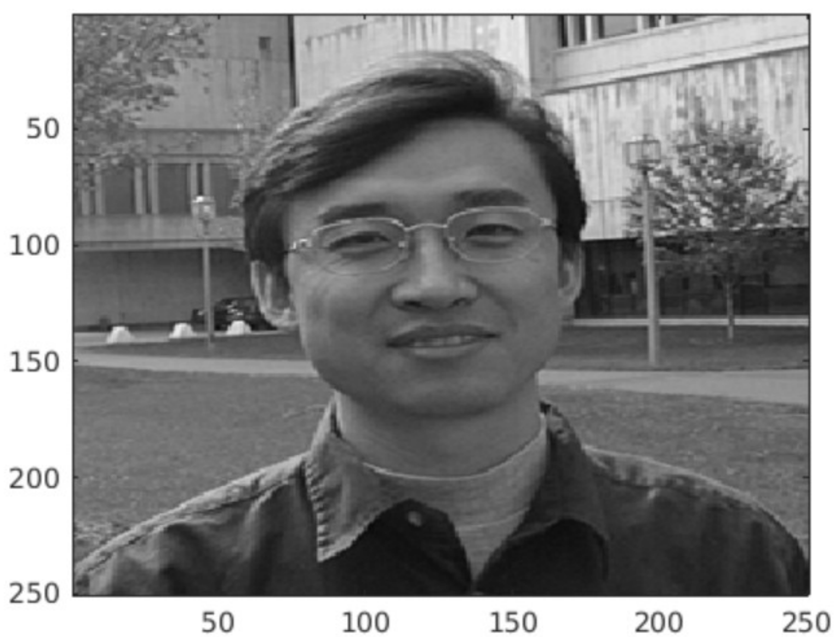
Programming Assignments

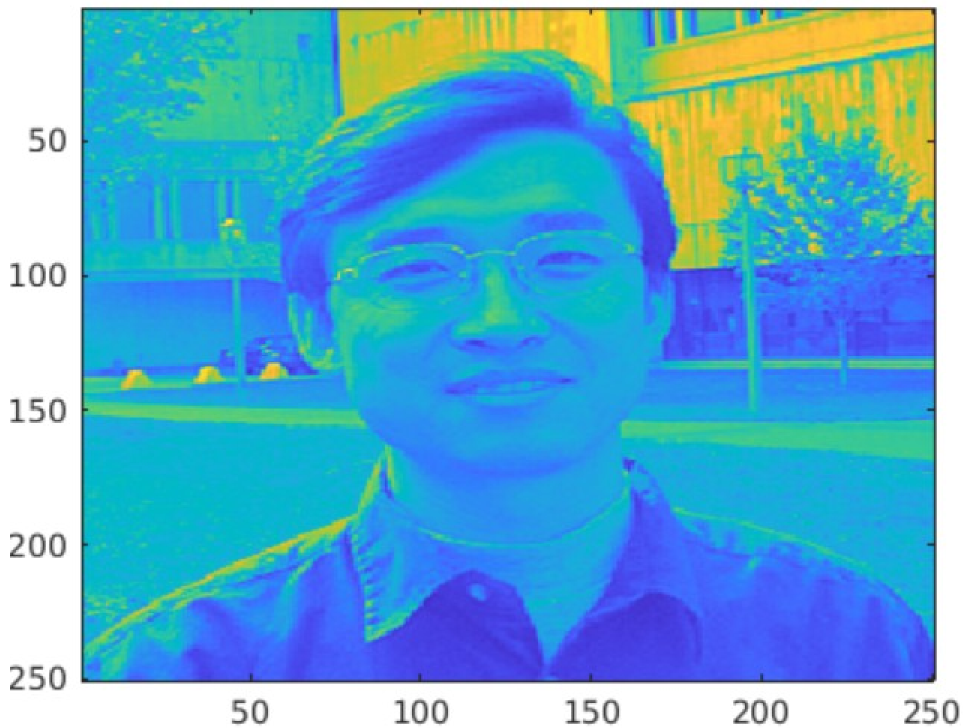
1. **Read in a color image C1(x,y) = (R(x,y), G(x,y), B(x,y)) in Windows BMP format, and display it.**



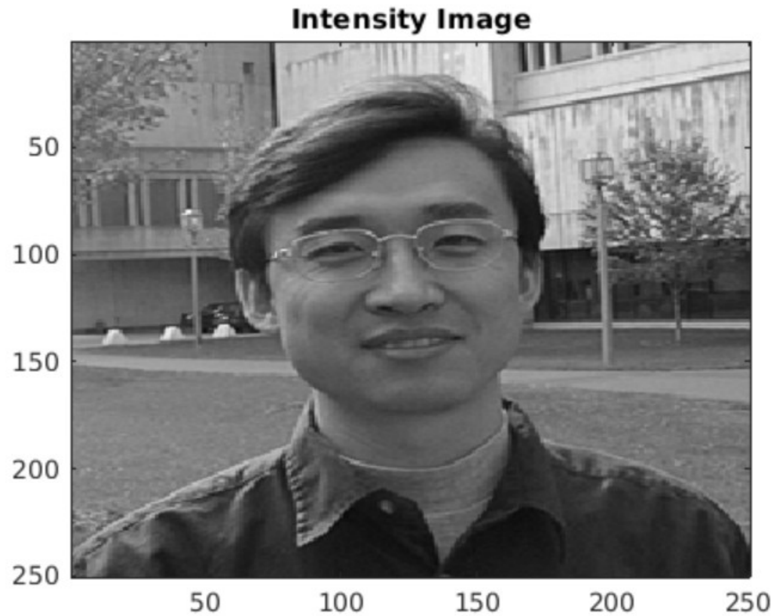
Display Color image in the RGB format and rest are red band, green band and blue band images

2. **Display the images of the three color components, R(x,y), G(x,y) and B(x,y), separately. You should display three black-white-like images.**



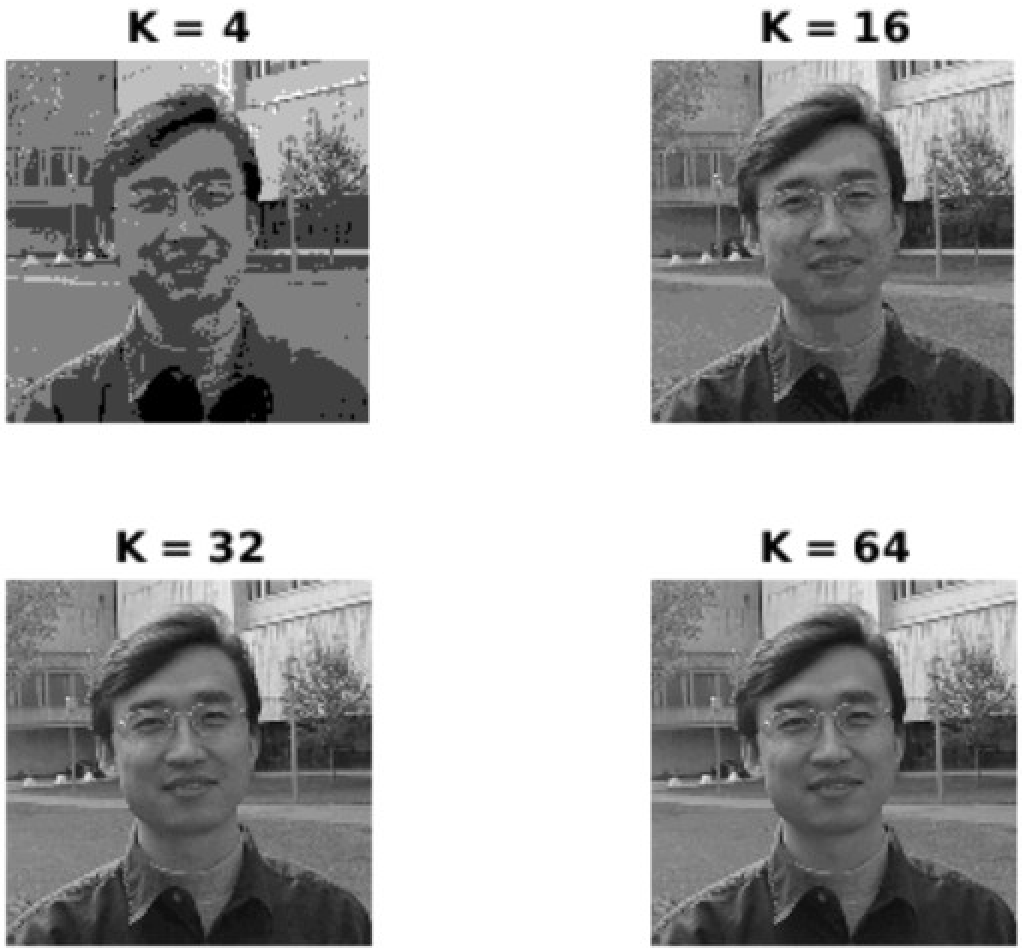


3. **Generate an intensity image I(x,y) and display it. You should use the equation $I = 0.299R + 0.587G + 0.114B$ (the NTSC standard for luminance) and tell us what are the differences of the intensity image thus generated from the one using a simple average of the R, G and B components.**



Green light is more sensitive then compared to red and least to blue color

4. **The original intensity image should have 256 gray levels. Please uniformly quantize this image into K levels (with K=4, 16, 32, 64). As an example, when K=2, pixels whose values are below 128 are turned to 0, otherwise to 255. Display the four quantized images with four different K levels and tell us how the images still look like the original ones.**



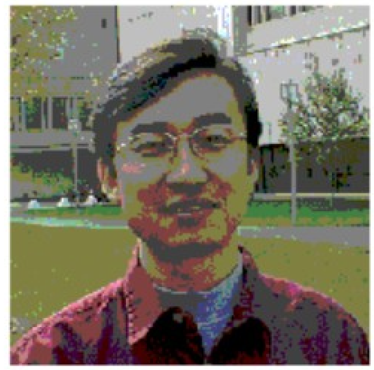
Since k=4 level and each level have 3 colors RGB so, 4^3 =64 which is very less and make image more darker and blur
With k=16 levels and each level have 3 colors RGB so 16^3 = 4096 image is okay,
With k=32 levels and each level have 3 colors RGB so 32^3 =32768 better than k=32,
With k=64 levels and each level have 3 colors RGB so 64^3 =262144 so,
we can conclude that as the k value increase image will be more clearly

5. Quantize the original three-band color image C1(x,y) into K level color images CK(x,y)= (R'(x,y), G'(x,y), B'(x,y)) (with uniform intervals) , and display them. You

may choose K=2 and 4 (for each band). Do they have any advantages in viewing and/or in computer processing (e.g. segmentation)?

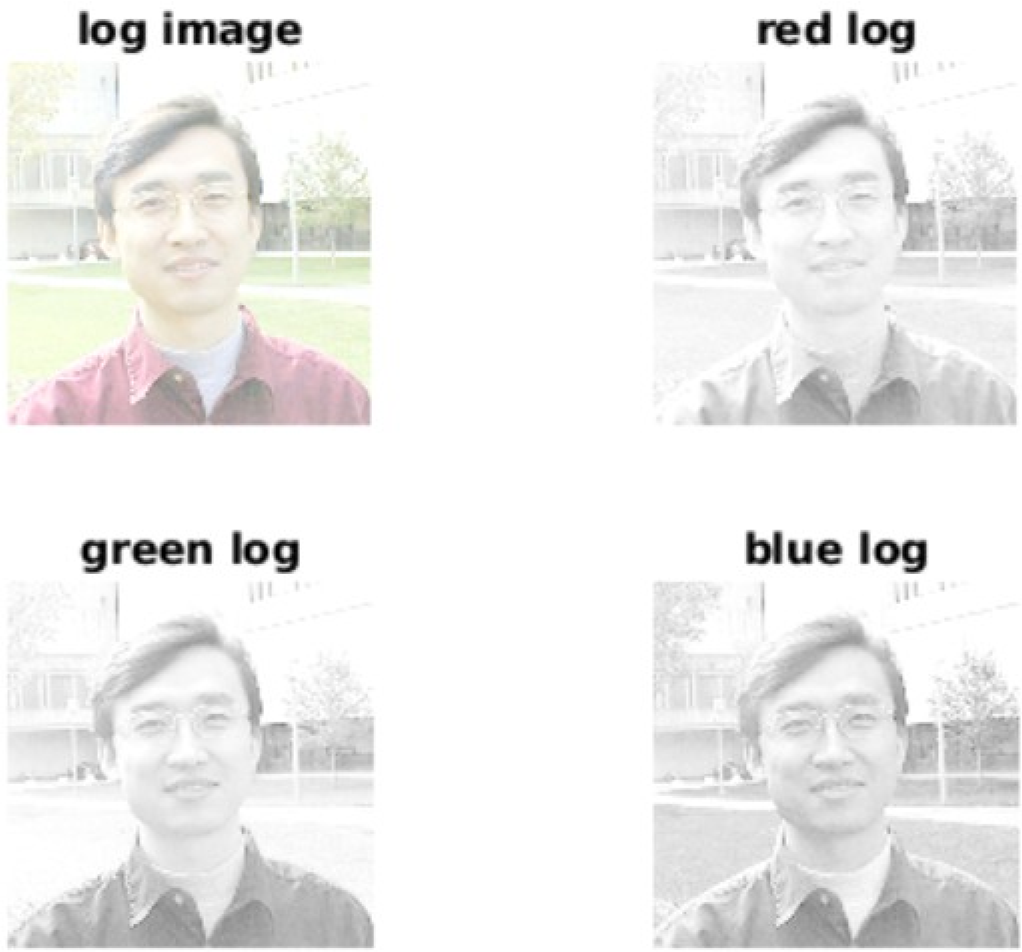
K = 2

K = 4



As the k level is very less the image is not clear it same with color image also.

6. **Quantize the original three-band color image $C1(x,y)$ into a color image $CL(x,y) = (R'(x,y), G'(x,y), B'(x,y))$ (with a logarithmic function) , and display it. You may choose a function $I' = C \ln(I+1)$ (for each band), where I is the original value (0~255) , I' is the quantized value, and C is a constant to scale I' into (0~255), and ln is the natural logarithm. Please describe how you find the best C value so for an input in the range of 0-255, the output range is still 0 – 255. Note that when I = 0, I' = 0 too.**



As the c value increase image fades, it not clear.