CS 5330 Week 3 Homework

Name: Hang Zhao
Email: zhao.hang1@northeastern.edu
Nuid: 002826538

Questions

1. Consider the situation where you want to put a camera in a car, looking towards the front of the car. A car is roughly 2.5m wide. You have a camera with a 1cm wide sensor with 2000 horizontal pixels.

A. If you want the car to be at least 50 pixels wide at 100m, what does your focal length need to be? Give the answer in mm. Be sure to identify which equation is relevant to the question before starting.

Answer:

we can use pinhole camera model equation, s = (f * W)/Z

- **s** is the size of the car in pixels (50 pixels in this case),
- **f** is the focal length (what we're solving for),
- **W** is the real-world width of the car (2.5m or 2500mm),
- **Z** is the distance to the car (100m or 100,000mm).

the camera sensor has 2000 pixels across 10mm of width. So, the pixel size is:

pixel size = sensor width / number of pixels = 10 / 2000 = 0.005mm/pixel.

we can change original equation s = (f * W)/Z, so that : f = (s*Z)/W*pixel size = (50*100000/2500)*0.005 = 100mm. so the focal length needs to be 100mm.

B. How many pixels would a car be in the image at 12.5m with the same focal length?

Answer:

s = (f*W)/(Z*pixel size) = (100*2500)/(12500*0.005) = 400 pixels.

the car will appear 400 pixels wide in the image.

C. Given a 10mm lens, do you need to make any adjustments for the scene to be in focus from 20-100m? (Clarification: you might need to make an adjustment if the change is bigger than 1 in 100.) Identify which equation is relevant to the problem before answering.

Answer:

we can use the thin lens equation:

$$\frac{1}{f} = \frac{1}{d_o} + \frac{1}{d_i}$$

- f is the focal length,
- d_0 is the object distance,
- d_i is the image distance.

$$d_i = \frac{1}{\frac{1}{f} - \frac{1}{d_o}}$$

when d_0 = 20m, d_i = 1/(0.1-0.00005) $\approx$ 10.005mm.

when d_0 = 100m, d_i = 1/(0.1-0.00001) $\approx$ 10.001mm.

the difference is 0.004mm less than 1% of 10mm. So there is no need to adjust the focus for distances between 20m and 100m with a 10mm lens because the change in focus distance is negligible (less than 1%).

2. Do an internet search to find information about the response distributions for human rods and cones.

A. What percentage of our cones are sensitive to long wavelength (red), medium wavelength (green), and short wavelength (blue)?

Answer:

- Long wavelength (Red-sensitive, L cones): About 60% of the cones.
- Medium wavelength (Green-sensitive, M cones): About 30% of the cones.
- Short wavelength (Blue-sensitive, S cones): About 10% of the cones.

B. When you look at the response curves for the red, green, and blue cones, how would you describe the differences between the three?

Answer:

The red and green cones dominate most of our color vision, with a lot of overlap, especially for yellow-green light. The blue cones are less noticeable, responsible for sensing blue light, but not so much involved in distinguishing brightness and detail. This is why we see blurry details in blue light, but are particularly sensitive to red and green light!

C. Given the information above, explain why the Bayer pattern on camera sensors has two green sensors for every one blue or red sensor.

Answer:

There are more green photosensitive units in the Bayer sensor because our eyes are more sensitive to green!

Think about it, in our retina, green cones (M cones) account for about 30%, while red cones (L cones) account for 60%. These two together almost dominate our perception of brightness, especially green. In contrast, blue cones (S cones) account for less than 10% and contribute much less to details.

Therefore, on the camera sensor, in order to imitate the visual characteristics of the human eye, the designer uses twice the green pixels, which can capture more brightness details while also meeting the visual focus of the human eye.

3. Do some research on Bayer Patterns / Bayer Filter (the wikipedia site is pretty good).

A. What are some issues that might arise because of the interpolation of colors to get RGB values at each pixel?

Answer:

When using Bayer Pattern/Bayer Filter, some problems may occur due to color interpolation. This is because each pixel actually only records one color (red, green or blue), and the camera needs to "guess" the complete RGB value through the color information of the surrounding pixels. The main problems are as follows:

1. Inaccurate color

Because the interpolation is calculated based on the color of the surrounding pixels, if the color difference between the pixels is large (such as the junction of red and blue), the camera may "guess" wrong, resulting in color distortion. For example, the red edge may appear to have a strange color mixture.

2. Aliasing

In some high-contrast scenes, such as when the edges of objects or textures are very fine, the interpolation algorithm may make the edges look "jagged" or a little blurry. For example, this problem will be particularly obvious when shooting fences or striped patterns.

3. Moire Pattern

When shooting fine repeating patterns (like fabrics, screens or building facades), the interpolation algorithm may introduce a strange "wave" or "rainbow" effect. This is because the arrangement and pattern of the Bayer Filter itself creates interference.

B. What is the benefit of saving RAW images and doing interpolation off the camera?

Answer:

1. More details and dynamic range: retain the complete information of bright and dark areas, less prone to overexposure or dead black.
2. Better color control: use more complex interpolation algorithms to get more natural and accurate colors.
3. Reduce camera limitations: computers can use more powerful algorithms, while cameras may be limited by hardware.
4. Avoid compression loss: RAW images are not compressed, retain the original data, and have higher image quality.
5. More flexible post-adjustment: you can freely adjust white balance, exposure, sharpening, etc. without losing image quality

C. Why might you want to use Cyan-Magenta-Yellow (CMY) instead of RGB filters? How would you get RGB values?

Answer:

The benefit of using a CMY filter is that it can capture more light and improve performance in low-light environments. Each RGB value can be obtained from CMY by subtracting.

R = 1 - C

G = 1 - M

B = 1 - Y

4. Color Spaces

A. When would it be important to use the CIE-Luv color space?

Answer:

The CIE-Luv color space is suitable for scenarios that require accurate processing of color perception, such as:

1. Color matching: Make the colors of displays and prints consistent.
2. Measuring color differences: It can better reflect the human eye's perception of color differences.

3. Uniform color distribution: When making heat maps or data visualization, color changes are more intuitive.

4. Image processing: Used to classify colors or analyze images, the results are more consistent with human eye perception.

B. For the YUV color space, the U channel is often called Blue - Yellow, and the V channel is often called Red - Cyan. Given the RGB to YUV conversion matrix, explain why U and V have those labels (you can find the matrix on Wikipedia or in my lecture notes).

Answer:

The YUV formula can be written in matrix form, which is roughly like this

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.436 \\ 0.615 & -0.515 & -0.100 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

Y is the brightness (Luma), which is obtained by mixing R, G, and B in proportion.

U and V are chroma (Chroma), which indicates the change of color.

U = - 0.147*R - 0.289*G + 0.436*B

As you can see, the U channel has a positive weight with blue (B), and a negative weight with red (R) and green (G).

This means that the U channel is used to capture the contrast between "blue" and "yellow":

Strong blue → increased U value.

Strong yellow (a mixture of red and green) → decreased U value.

Therefore, the U channel emphasizes the difference between blue and yellow, and is therefore called the blue-yellow channel.

V = 0.615*R - 0.515*G - 0.100*B

Here, the V channel and red (R) have positive weights, while green (G) and blue (B) have negative weights.

This means that the V channel is used to capture the contrast between "red" and "cyan":

Strong red → increased V value.

Strong cyan (a mixture of green and blue) → decreased V value.

Therefore, the V channel is called the red-cyan channel.

C. Is there any connection between the UV definitions and the human visual system?

Answer:

The connection between the YUV color space and the human visual system is that it cleverly utilizes our sensitivity to brightness and color. The Y channel (brightness) focuses on light and dark information because we are most sensitive to changes in brightness, which ensures the clarity of the image; while the U and V channels (chrominance) capture the contrast between blue-yellow and red-cyan, which is consistent with our perception of color, but their data can be compressed because we are not as sensitive to color details as to brightness. This design not only conforms to the characteristics of the human eye, but also can efficiently store and process images, which is very suitable for video and image compression technology.