Name: Harsh Vijay Mamania

CS 5330: Week 3 Homework

Question #1

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

Solution-

Here's when CIE-Luv color space would be particularly important:

Perceptual uniformity requirements

- Provides perceptually uniform color representation where equal Euclidean distances correspond to approximately equal perceived color differences by human observers
- Valuable for color difference metrics and quality control applications
- Essential when quantifying perceived color differences is required

Additive color systems and display applications

- Specifically designed for additive color mixing systems
- Well-suited for self-luminous displays such as monitors, projectors, and LED screens
- Performs optimally with emissive color systems rather than reflective surfaces

Separation of luminance and chromaticity

- Enables independent manipulation of lightness (L*) and chromaticity (u*, v*) components
- Particularly useful for color correction, color grading, and white balance adjustments
- Cylindrical representation (LCh_uv) allows intuitive control of hue and saturation parameters

Applications requiring device-independent perceptual color

- Provides standardized, perceptually meaningful color space unlike device-dependent RGB
- RGB lacks perceptual uniformity, making CIE-Luv preferable for consistent color perception
- Important for tasks requiring consistent color perception across different viewing conditions

Screen-based workflows

- Generally preferred over CIE-Lab for display-oriented applications
- CIE-Lab is better suited for reflective surfaces and subtractive color mixing (e.g., printing)

- CIE-Luv is the choice for screen-based work, while CIE-Lab is preferred for print/physical materials

---

Question #2

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).

Solution-

Based on the conversion matrix from Professor Bruce's lecture notes:

U = -0.30R - 0.59G + 0.38B
V = 0.58R - 0.59G - 0.11B

- **U channel as "Yellow - Blue"**
    - The U channel has a strong positive coefficient for Blue (+0.38) and negative coefficients for Red (-0.30) and Green (-0.59)
    - When Blue is high and Red/Green are low, U becomes strongly positive (Blue dominates)
    - When Red and Green are both high but Blue is low, this produces Yellow (R+G=Yellow), making U strongly negative
    - Therefore, U represents a spectrum from Yellow (negative values) to Blue (positive values)

- **V channel as "Red - Cyan"**
    - The V channel has a strong positive coefficient for Red (+0.58) and negative coefficients for Green (-0.59) and Blue (-0.11)
    - When Red is high and Green/Blue are low, V becomes strongly positive (Red dominates)
    - When Green and Blue are both high but Red is low, this produces Cyan (G+B=Cyan), making V strongly negative
    - Therefore, V represents a spectrum from Cyan (negative values) to Red (positive values)

- **Opponent color encoding**
    - These channels encode color in an opponent manner, similar to human visual perception
    - Each chrominance channel represents a pair of complementary/opponent colors
    - The sign (positive/negative) indicates which color in the pair dominates
    - The magnitude indicates the intensity or saturation of that color component

---

Question #3

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

Solution-

Connection between the U and V channels and the human visual system:

Opponent color theory alignment

- The UV color channels are based on opponent color theory, which models how the human visual system processes color information
- Human vision processes color through opponent channels: Red-Green, Blue-Yellow, and Light-Dark (luminance)
- The U channel (Yellow-Blue) and V channel (Red-Cyan, which approximates Red-Green) directly reflect this biological color processing mechanism

Neural processing in the retina and LGN

- Cone cells in the human retina (L, M, S cones) send signals to ganglion cells that compute opponent color differences
- Some ganglion cells compute (L+M) - S responses (roughly Yellow-Blue), corresponding to the U channel
- Other ganglion cells compute L - M responses (roughly Red-Green), corresponding to the V channel
- This opponent processing continues through the Lateral Geniculate Nucleus (LGN) to the visual cortex

Separation of luminance and chrominance

- The human visual system separates brightness information (Y channel) from color information (U and V channels)
- This separation is efficient because humans are more sensitive to luminance changes than chrominance changes
- The Y channel mimics the luminance sensitivity of human vision by weighting Green most heavily (0.59), matching the eye's peak sensitivity to green wavelengths

Perceptual efficiency and compression

- Because UV channels align with human color perception, they enable efficient video compression
- Human vision has lower spatial resolution for color (chrominance) than brightness (luminance)
- YUV encoding exploits this by allowing chrominance subsampling without perceptible quality loss
- This mimics the distribution of photoreceptors in the eye (more luminance-sensitive rods than color-sensitive cones in peripheral vision)

Question #4

I am building a convolutional network. Given an image with 3 channels (RGB), I want the first convolution layer to have 16 5x5 filters. How many filter coefficients (parameters) does the layer need to learn? Explain your reasoning.

Solution-

- Each filter must span all input channels.
  - Input has 3 channels (R, G, B).
  - A convolution filter in the first layer is therefore $5×5×3$, not just $5×5$.

- Parameters per filter (weights):
  - $5×5×3=75$ coefficients.

- Total parameters for 16 filters (weights only):
  - $16×75=1200$ coefficients.

- If the layer also has a bias per filter (common in Conv layers):
  - Add 16 bias terms.
  - Total = $1200+16=1216$.

Answer:
- **1200** learnable filter coefficients (weights), or **1216** total parameters if you include one bias per filter.

Question #5

Given a second convolution layer on the same network (first layer has 16 filters), we want the second layer to have 32 3x3 filters. How many filter coefficients (parameters) does the second layer need to learn? Explain your reasoning.

Solution-

- The input to the second convolution layer is the output of the first layer.
  - The first layer has 16 filters, so it produces 16 feature maps (channels).

- Each filter in the second layer must span all input feature maps.
  - Therefore, each filter has size $3 \times 3 \times 16$, not just 3×3.

- Parameters per filter (weights):
  - $3 \times 3 \times 16 = 144$ coefficients.

- Total parameters for 32 filters (weights only):
  - $32 \times 144 = 4608$ coefficients.

- If the layer includes a bias per filter (common in practice):
  - Add 32 bias terms.
  - Total = $4608 + 32 = 4640$.

Answer:

- **4608** learnable filter coefficients (weights), or **4640** total parameters if biases are included.

Question #6

When building a histogram, what are the parameter decisions you need to make?

Solution-

Number of bins

- Determines how many discrete levels to quantize the pixel intensity or feature values into
- For grayscale images: Common choices are 256 bins (one per intensity level 0-255), or fewer bins like 64, 32, or 16 for coarser quantization
- For color histograms: Must decide bins per channel (e.g., $8 \times 8 \times 8 = 512$ bins for RGB)
- Trade-off: More bins capture finer detail but are more sensitive to noise and illumination changes

Range of values

- Define the minimum and maximum values to include in the histogram
- For standard 8-bit images: typically [0, 255]
- May need to adjust for normalized images [0, 1] or other representations
- Can crop or clip extreme values if dealing with outliers

Which channels or features to use

- Single channel: Grayscale intensity, or individual R, G, B channels separately
- Multiple channels: Joint histogram across RGB, HSV, or other color spaces
- Feature-based: Histograms of gradients (HOG), orientations, or other computed features

Spatial considerations

- Global histogram: Compute over entire image
- Local/regional histograms: Divide image into blocks or regions
- Spatial binning: Include location information alongside intensity/color

Normalization

- Whether to normalize histogram by total pixel count (convert to probability distribution)
- Useful for comparing images of different sizes or for histogram matching applications

---

Question #7

(Can be done this week or next) Take a picture of a cast shadow outside on a sunny day, ideally on a sidewalk, asphalt, or some other surface that is close to grey. Use my 2-D histogram code (or your own code) to make an rg chromaticity histogram of the image. If you want a more physically correct result, take the image in RAW format (most devices have a way you can do this) and use this code.Download this codeDownload this code to convert it to a 16-bit TIFF that can be read by OpenCV. You will need the RawPy Python library, which can be installed using pip.

Taking RAW Images with Mobile Devices:
https://blackeyelens.com/why-you-should-shoot-in-raw-with-mobile/

RawPy library: https://pypi.org/project/rawpy/

If you took a picture that includes only the neutral material with a shadow, you should see a white/yellow peak (lit area) and a bluish peak (shadowed area) in the histogram. Include the picture and the histogram in your pdf.

Solution-

<I will complete this assignment next week, as per the provided option in the instructions.
Due to the snowstorm in Boston this weekend, there was insufficient sunlight to capture a suitable image with clear cast shadows on a neutral grey surface.
I plan to take the photograph on a sunny day next week and generate the rg chromaticity histogram.>
