# Introduction to Linux VisionX and Image presentation

## Lab 1 Questions

### A. General Image Manipulation
1.  What is the pixel coordinates of Albert's right eye?
    The pixel coordinate is (106, 154.0)

2.  How did you make this measurement?
    I made the pixel coordinate measurement by hovering my cursor over the eye. Then, the pixel coordinate is displayed in green font on the bottom-left corner of the image.

3.  What is the width of Albert's mustache?
    The width is 47.07mm. This is measured by using "line" measurement.

4.  What are the main issues in making distance measurements on images? Consider and comment on the following:

    Mouse control: There is no problem with the mouse control. However, we would have to go back-and-forth between the measure tool and navigation tool in order to measure as accurate as possible.

    Feature visibility: It is difficult to tell the boundary of the moustache. I tried zooming in into the image, and it is still difficult to determine the boundaries. So, it is difficult to determine start and end points.

    Image Size: I did not find image size to be a big problem. However, we have to zoom in to get a closer view of the moustache and to try and determine the parts of the moustache.

    Question ambiguity: I find the question ambiguous in asking to measure Albert's moustache. This is because the points in which the measurement should start and end is not clearly defined. We could have measured the width using the "line" measurement a bit to the top or bottom, and the result may be different to a few millimeters.

### D. Image Set Display, Pixel Quantization
1.  What does the girl image sequence show?
    It shows how the changes in the number of total colors in the picture affects the picture.

2.  What do the image parameters printed in the caption of the image mean?
    The "grey" is the number of greyscale colors in the picture, or also called as grey levels. The "bpp" is the bits per pixel. For example, it there are 3 bpp, then the maximum number of colors are 8 ($2^3$).

3. At what quantization does degradation in image quality become noticeable?
It became noticeable when it reaches bits per pixel of 4 with grey level of 10, and then when the number of colors go further down, the degradation became very clear.

4. What is the implication of this observation for image display design? Comment on the number of gray levels you would include in a product for general use and why?
This observation tells us the relation of the bits per pixel and grey level with the overall quality of the image. We can see how it is difficult to notice the difference in quality between bpp of 8 to bpp of 5. With this knowledge, we can determine whether it is worth it to have higher bpp in expense of memory size of the image. I would use the grey level of 64, which is of bpp of 6. This is because we can keep the image to be in a high quality to view, but to also keep the memory size of the image to be smaller.

5. Many home theater systems claim to have a 4000 to 1 contrast ratio or (much) greater. Assuming that this is possible by the image encoding methods used how is such a large range useful given human visual system limitations?
Human visual system's static contrast ratio is certainly below the ratio of 4000:1, as a result the contrast ratio of home theater system may sound like an overkill. However, human also has a dynamic range or dynamic contrast ratio limit. So it may take some time for human to adapt to the change in contrast, but will then be able to fully view it. As a result, the 4000:1 contrast ratio will actually be useful for humans as they are fully able to view the whole range with a bit of adaptation, and will allow human to enjoy both the darkest and brightest colors.

## E. Terminal command and log enhancement

1. What does the vmath command do? The vmath with compute a mathematical function of the pixels in the VisionX image passed into the command. For our case, we did "vmath -mlog -scale girl-grey.vx of=girl-log.vx". The "-mlog" flag means to pass a modified log function, which is a log function but values below zero are set to zero and values between 1-3 are linearly scaled. The "-scale" flag will set the range of the output pixel to be from 0 to 255, which is the default value. The "girl-grey.vx" is the input to the command and "of=girl-log.vx" is the output file, which is generated by the command.

2. How does the vmath operation affect the visibility of image features?
Are some parts of the image easier to see after the log transform?
Is it possible to see more or less detail after the transform?
The vmath operation that we use, which is a log function, makes the lower range of pixel value to be easily differentiated with each other, as they scaled to be significantly different between each other. As a result, the darker pixels are now more visible and the boundaries between them is easily seen. On the other hand, because of the log function, the originally higher pixels become less differentiable between each other, as their values become closer to each other due to the log function.

Yes, some parts of the image are now easier to see after the log transform, especially in parts that are originally darker. An example is how it is easier to differentiate and draw boundaries between the pupil and iris of the person.
In terms on whether there are more or less detail after transformation, I would say that we can now see more details. After the transformation, we can immediately see the details on the eyes, leaves, and the clothing that the woman is wearing.

3. If more detail is visible, how is this possible? (consider your answer to question 2) (think carefully)
The increase in detail is because of the log function that is used to transform the image. In log function, the darker or lower values pixels are expanded more through the pixel range. On the other hand, the brighter pixels are more compressed to the higher range of pixels. As a result, it is now easier to see the details of the of the parts of the image that is darker, as the pixels between them are expanded. This makes dark boundaries to be clearer. One example is the eyes, in which it is easier to compare the pupil and iris. Another example is the leaf, which is originally dark, but with this enhancement we are able to see the boundaries of the leaf.

## F. Large Image Display

1. What is the size of the x-ray image?
The size is 2046mm x 2046mm.

2. What are the window settings?
The window value is set to 5. The level is kept at default at 2049. The window is set to be very small to be able to clearly see the effects of interpolation. In order to view the high contrast, window value 100 is enough, but it is changed to 5 in analyzing the performance of interpolation.

3. What is the range of pixels in the image?
The range of the pixels is from 1586 to 3235.

4. Which interpolation method is the best and why?
Through my experimentation, the bicubic interpolation is the best, compared to bilinear or no interpolation. Having no interpolation is the worse as the boundaries are too sharp due to the pixelation. The bicubic is better than the bilinear, although not by much. We can see that bicubic interpolation is sharper and is better visually. It also has a clearer boundary between the dark and bright pixels.

## G. Segmentation Using Thresholding

1. What is the best threshold for the facsimile image?
The best threshold is 211.

2. What is your criterion for best threshold?
The best one would be the one that divides the background and foreground. In this case, it would be the one that will allow most visibility of the writing on top and middle of the image, and the figure on the bottom. If the threshold is too low, then some of the letters may be disappearing. If threshold is too high, then there is too much noise of black dots around the letters and lines, that disturbs the visibility.

3. What is the best threshold for the map image?
   The best threshold is 96.

4. What is your criterion for best threshold?
   The best one is the one that still has a clear visibility of almost the entire map, while also having the street name labels to be visible enough for human to read.

5. What is the problem in thresholding the map image?
   The main problem in thresholding this image is that the outer parts of the image is darker (lower pixel values) compared to the middle part of the image. As a result, if we go for a higher threshold value, then the darker outer parts of the map may not be visible as they become totally dark, although we have a better visibility of the street names in the middle of the map. However, if we go for a lower threshold, then the street names may barely be visible, which may make the map to be pointless. As a result, there needs to be a balance and sacrifice on whether we prefer for a larger map or a perfect visibility of the street names.

6. How might you get a better result on the map image? (That is, what kind of processing operations might improve the result?)
   One way that I thought about is to divide the image into different parts. This allows us to separate the darker outer parts and the brighter middle parts. Then, we apply different threshold values for different parts to provide the best visibility of the map. Then, we combine the parts together again and we will have a better overall map image.
   Increasing the window value seemed to help a bit as some parts may become clearer. However, it makes the image to be darker, which will not be optimal.

## H. Using vview and vdview, Edge Detection

1. Are all the edges detected by vedge?
   No, not all the edges on the image is detected by vedge. We can clearly see the outline of the caboose. However, not all the edges are detected. For example, not all the edges on the railway is detected.

2. Can you improve the result? (by using different options)
   Yes, the result can be improved. The first method that I tried is using vedge, but with the flag –g which will use gradient instead of edge output. When I set the window to 1 and level to 30, I was able to detect all of the objects (caboose, railway, and fence), without too much noise from the background. The second method that I tried is using vsobel command, which is a sobel edge detection function. When setting window to 1 and level to 12, I was again able to detect all of the objects. Vsobel seemed to provide a better result as it provides a more detailed and refined edges.

3. What is an edge? (a) provide a definition for an edge:
   An edge is a set of connected pixels where there is a sharp and abrupt change in brightness or intensity. In Image Processing, edges are often used to indicate the outline of objects.

4. Describe what you mean by edges in the context of the caboose image.
   In the context of the caboose image, the edges, which would be pixels outlining the drastic change in intensity, would be the outline or boundary of the objects in the image, which will separate the objects with each other or the background (which is the wall). So, this edges will be outlining the different objects, such as the caboose, the railway, and the fence.