EE5470 Computer Vision Lab 1

NAME: Xinyi(Blanca) Bai ID: xb52

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?

(130,152,0) (142,160,0) (148,160,0) (155,156,0) (146,152,0) (142,152,0)

1. How did you make this measurement?

I choose this six points as the key points to describe his right eyes:



1. What is the width of Albert’s mustache?
2. What are the main issues in making distance measurements on images? Consider and comment on the following:

Mouse control: First select “Line” in the “Measure”, then click on the Mouse at the left corner of Albert’s mustache, then drag mouse to the right corner of Albert’s mustache(make sure the value from the message area of this right corner is 255), then click “Do Measure”, so the width will be shown in the top left corner of the image, which is that green-color words with black ground. From my measurement, it is “Dist:47.5”.

Feature visibility: It is hard to figure out the left corner of Albert’s mustache.

Image Size:

256\*256

From the Image Statistics for ein.vs

Question ambiguity:

what does width mean? To measure the width, should we contain the lips near mustache? What does image size mean? Does it means how many storage this image take to be stored in our computer? Or how many pixels are in this image?

**D. Image Set Display, Pixel Quantization**

1. What does the girl image sequence show?

The image become darker and darker, noise become more and more, and image seems lack exposure, the numbers of colors become less and less. In the message area, the parameter “grey” decrease from 256 to 2 gradually, and “bpp” decrease from 8 to 1 gradually.

1. What do the image parameters printed in the caption of the image mean?

“grey” means “grey value” which indicates the brightness of a pixel. The minimum grey level is 0. The maximum grey level normally is 255.

“bpp” means bit per pixel, it can also explains as numbers of different colors, for example:

|  |  |
| --- | --- |
| **Bits per pixel** | **Number of colors** |
| 1 bpp | 2 colors |
| 2 bpp | 4 colors |
| 3 bpp | 8 colors |
| 4 bpp | 16 colors |
| 5 bpp | 32 colors |
| 6 bpp | 64 colors |
| 7 bpp | 128 colors |
| 8 bpp | 256 colors |
| 10 bpp | 1024 colors |
| 16 bpp | 65536 colors |
| 24 bpp | 16777216 colors (16.7 million colors) |
| 32 bpp | 4294967296 colors (4294 million colors) |

1. At what quantization does degradation in image quality become noticeable?

when “nav” is 9, when “grey” is 9, “bpp” is 4, the image does degradation remarkably.

1. 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?

The grey level influence the exposure of a image, the less the grey level is, the darker we will feel, and the more noise will appear. But when the grey level become very small, we suddenly feel the noise becomes less, maybe the noise gradually become part of the dark or light background since the dark area or light area become larger and larger.

I think the grey level = 21 is the minimum level in a product for general use, because if less then 21, there will be many noises and these noise are easy to be recognized, so it gives us bad visual experience.

1. 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?

Contrast ratio is the ratio of the luminance of the brightest color (white) to that of the darkest color (black).

The retina of human eye has a static contrast ratio of around 100:1, so over 100:1, we are unable to perceive a contrast ratio of 1000:1, over 100:1, our human eyes cannot tell the difference.

**E. Terminal command and log enhancement**

1. What does the vmath command do?

Compute a primitive mathematical function of the pixels of a VisionX image.

1. 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?

After we use vmath -log to compute the logarithm of all pixel values, the image become very bright, white.

We can also use vmath -abs to compute the absolute value of all pixels, vmath -exp to compute the exponential of all pixel values.

1. If more detail is visible, how is this possible? (consider your answer to question 2) (think carefully)

According to the tutorial in [www.via.cornell.edu/visionx/cguide/fpage.html](http://www.via.cornell.edu/visionx/cguide/fpage.html)**.**

More operation can be used combined with vmath. Such as

|  |  |
| --- | --- |
| bf=bval | The value of the parameter bval is added to all pixels |
| tf=tval | all pixels are multiplied by the parameter tval. |
| hi=hval | the high value for the pixels in the output file is set to hval. The default value for hval is 1.0 for all formats except unsigned byte (format 1) in which case the default is 255. |
| -scale | set the range of the output pixels from lval to hval. |
| -clip | set all pixels with values higher than hval to hval and set all pixels with values less than lval to lval. |

**F. Large Image Display**

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

2048\*2048

From the Image Statistics for xray.vs

1. What are the window settings?

Window: 800

Level: 2049

1. What is the range of pixels in the image?

1-4096

From the Image Statistics for xray.vs: maximun pixel value is 4096. Minimun pixel value is 1.

1. Which interpolation method is the best and why?

I think bicubic is the best, because bicubic method is more clear, smooth than bilinear interpolation, and bicubic interpolation contains less artifacts.

Then I searched the two methods in the internet, It is said “bicubic interpolation makes use of more data, its results are generally smoother. Bicubic interpolation creates smoother curves than bilinear interpolation”

**G. Segmentation Using Thresholding**

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

level = 213

1. What is your criterion for best threshold?

My criterion is to make every English letter on the top of the image as clear as possible, and make the small hole in the letter ‘e’ still a hole, instead of being filled with black colors.

1. What is the best threshold for the map image?

Level = 93

1. What is your criterion for best threshold?

Don’t let the black dot cover any name of the location in the left part of the image, and make the name of the location in the right part as clear as possible.

1. What is the problem in thresholding the map image?

When the name of location in the right part of the image is clear to be read, the name of location in the left part of the image is covered by a lot of black dot and unable to be read.

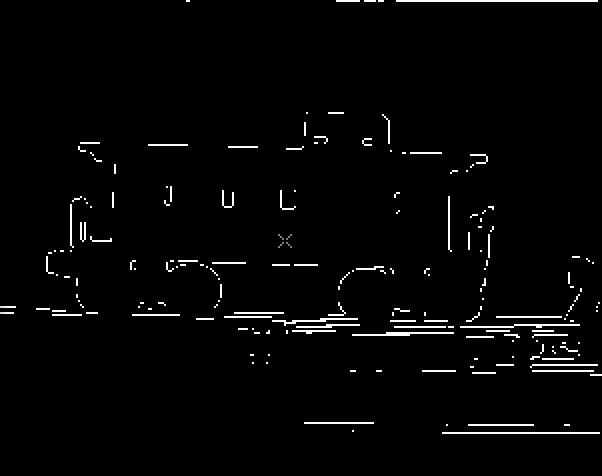
1. How might you get a better result on the map image? (That is, what kind of processing operations might improve the result?)

Increase the Window value, maybe set it as 256, and then change level value to see what will happen.

**H. Using vview and vdview, Edge Detection**

1. Are all the edges detected by vedge?

Not all edges are detected. Only some edges,



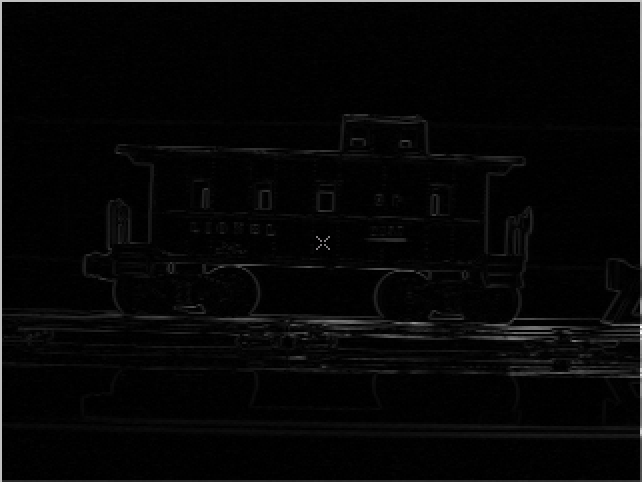
1. Can you improve the result? (by using different options)

I tried the other edge detection algorithm. The result of vderiche method is the best. All results are as follows:

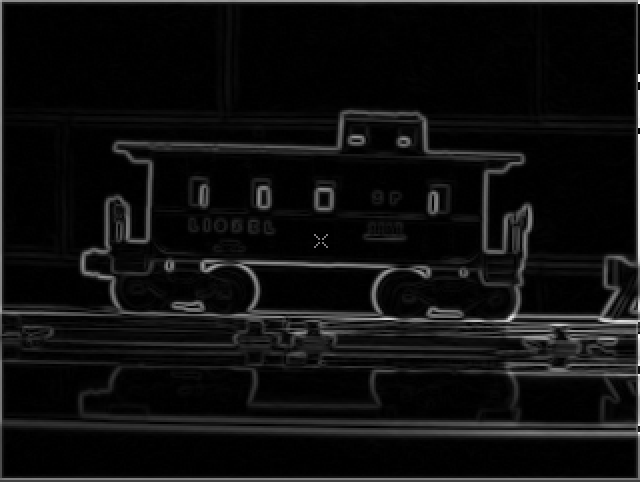
The result of vsobel method:



The result of vrcross method



The result of vderiche method



The result of vedgex

I cannot implement vedgex methof on this image, because vedges method needs the input image be two channel file, while this image is not two channel image.

1. What is an edge? (a) provide a definition for an edge:

An edge is a boundary or contour of the important part (object) in an image. What is important? It depends on your research focus.

1. Describe what you mean by edges in the context of the caboose image.

So under my definition, the target of this caboose image is the car/train, so the edge is the contour of the car and it’s windows.