Computer Vision HW1 Report

B04507009 電機四 何吉瑞

Code usage:

Conventional:

python3 rgb2gray.py --source_path [the path of bgr image folder] --target_path [the folder path to save grayscale images] --img_name [the filename of .png image]

Advanced:

python3 adv_rgb2gray.py --source_path [the path of bgr image folder] --target_path [the folder path to save grayscale images] --img_name [the filename of .png image] --already_done [0 for rerunning and 1 for testing] --use_ans [1 for using opencv_contrib] --r_factor [the ratio of radius to sigma_color] --save_img [save all images] --npy_path [the path of .npy files about the saved results]

For testing, it's only required to adjust the black arguments.
The results are already saved in .npy files in npy_path, which needs not to be changed if the code is compiled in original folder "hw1/src/".
If it is needed to rerun all the BJF of candidate images, please change the argument "already_done" to 1.

Example command (only load the answer):
python3 adv_rgb2gray.py --source_path ../testdata --target_path ../result --img_name 0a.png

1. Conventional RGB2GRAY

After rerunning the code, a grayscale image will be generated in target_path. I use Y = B + G + R for conversion. The reference is the function cvtColor() in OpenCV. After testing, the results of my conversion and that in OpenCV are same. An example is shown in Fig 1.

(a) (b) (c)

Fig 1. (a) Original BGR image. (b) Grayscale image from my conversion. (c) Grayscale image from the conversion of OpenCV.

2. Advanced RGB2GRAY

(1) Rerunning description:

After rerunning the code, the following files and folders will be generated in target_path (for simplicity, I call it “result/”):

result/can/

The grayscale candidate images, which are the answers.

result/res/

Results of using candidate images as guided images for joint bilateral filter.

result/surface/

Nine 2D surfaces for the cost of each selected parameter.

result/vote.txt

The voting results.

The answer images for homework is saved in result/can/. Note that the index of image is larger, the votes of image is higher.

(2) Tied ballot:

I generate all candidate images with tied ballot. For example, if the voted results are [4, 4, 3, 3, 2, …]. The code will generate four candidate images.

(3) Verification:

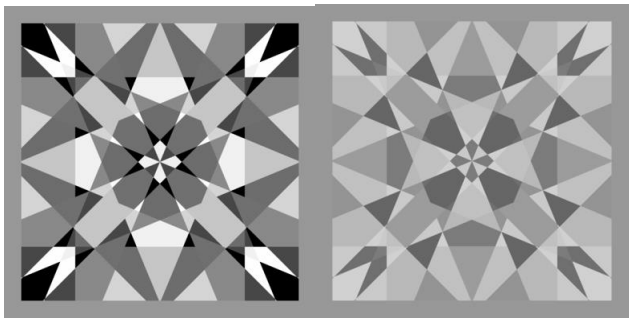
I check the filtered images with the result from function in OpenCV_contrib. The pixel wise difference is at most 0.4, which I think is small enough. The difference may because of the contour in OpenCV is a circle instead of a square.

(4) Results

0a

Weights of BGR	votes	filename
Y = 0.0B + 0.0G + 1.0R	9	0a_y0.png
Y = 1.0B + 0.0G + 0.0R	9	0a_y1.png

Fig. 2 The candidate images of 0a



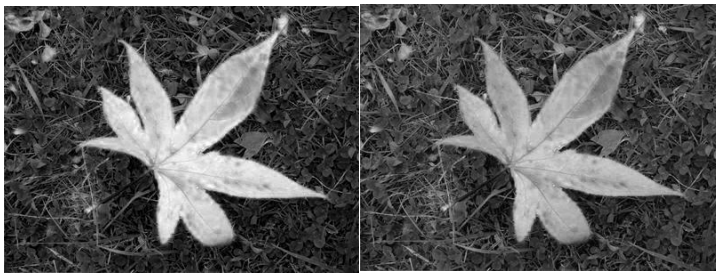
0a_y0.png 0a_y1.png

0b

Weights of BGR	votes	filename
Y = 0.0B + 1.0G + 0.0R	2	0b_y0.png
Y = 0.1B + 0.2G + 0.7R	2	0b_y1.png
Y = 0.0B + 0.0G + 1.0R	3	0b_y2.png
Y = 0.0B + 0.2G + 0.8R	3	0b_y3.png

Fig. 3 The candidate images of 0b

0b_y0.png 0b_y1.png



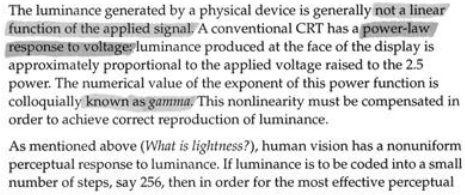
0b_y2.png 0b_y3.png

0c

Weights of BGR	votes	filename
Y = 0.3B + 0.4G + 0.3R	3	0c_y0.png
Y = 0.0B + 0.0G + 1.0R	2	0c_y1.png
Y = 0.3B + 0.5G + 0.2R	1	0c_y2.png
Y = 0.3B + 0.3G + 0.4R		0c_y3.png
Y = 0.2B + 0.6G + 0.2R		0c_y4.png

Y = 0.2B + 0.5G + 0.3R		0c_y5.png
Y = 0.1B + 0.5G + 0.4R		0c_y6.png

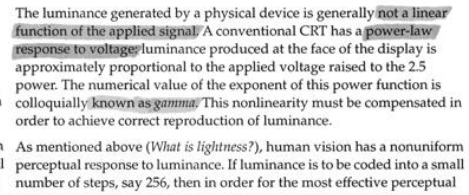
Fig. 3 The candidate images of 0c



The luminance generated by a physical device is generally not a linear function of the applied signal. A conventional CRT has a power-law response to voltage: luminance produced at the face of the display is approximately proportional to the applied voltage raised to the 2.5 power. The numerical value of the exponent of this power function is colloquially known as *gamma*. This nonlinearity must be compensated in order to achieve correct reproduction of luminance.

As mentioned above (*What is lightness?*), human vision has a nonuniform perceptual response to luminance. If luminance is to be coded into a small number of steps, say 256, then in order for the most effective perceptual

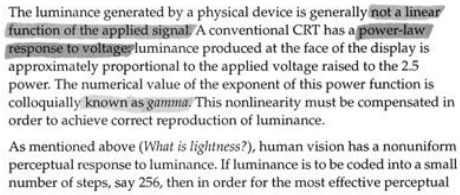
0c_y0.png



The luminance generated by a physical device is generally not a linear function of the applied signal. A conventional CRT has a power-law response to voltage: luminance produced at the face of the display is approximately proportional to the applied voltage raised to the 2.5 power. The numerical value of the exponent of this power function is colloquially known as *gamma*. This nonlinearity must be compensated in order to achieve correct reproduction of luminance.

As mentioned above (*What is lightness?*), human vision has a nonuniform perceptual response to luminance. If luminance is to be coded into a small number of steps, say 256, then in order for the most effective perceptual

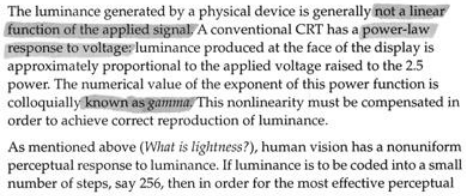
0c_y1.png



The luminance generated by a physical device is generally not a linear function of the applied signal. A conventional CRT has a power-law response to voltage: luminance produced at the face of the display is approximately proportional to the applied voltage raised to the 2.5 power. The numerical value of the exponent of this power function is colloquially known as *gamma*. This nonlinearity must be compensated in order to achieve correct reproduction of luminance.

As mentioned above (*What is lightness?*), human vision has a nonuniform perceptual response to luminance. If luminance is to be coded into a small number of steps, say 256, then in order for the most effective perceptual

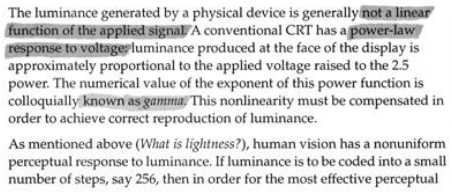
0c_y2.png



The luminance generated by a physical device is generally not a linear function of the applied signal. A conventional CRT has a power-law response to voltage: luminance produced at the face of the display is approximately proportional to the applied voltage raised to the 2.5 power. The numerical value of the exponent of this power function is colloquially known as *gamma*. This nonlinearity must be compensated in order to achieve correct reproduction of luminance.

As mentioned above (*What is lightness?*), human vision has a nonuniform perceptual response to luminance. If luminance is to be coded into a small number of steps, say 256, then in order for the most effective perceptual

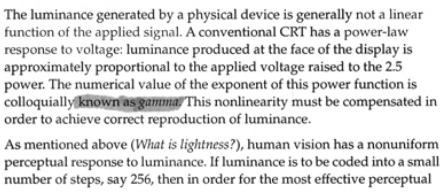
0c_y3.png



The luminance generated by a physical device is generally not a linear function of the applied signal. A conventional CRT has a power-law response to voltage: luminance produced at the face of the display is approximately proportional to the applied voltage raised to the 2.5 power. The numerical value of the exponent of this power function is colloquially known as *gamma*. This nonlinearity must be compensated in order to achieve correct reproduction of luminance.

As mentioned above (*What is lightness?*), human vision has a nonuniform perceptual response to luminance. If luminance is to be coded into a small number of steps, say 256, then in order for the most effective perceptual

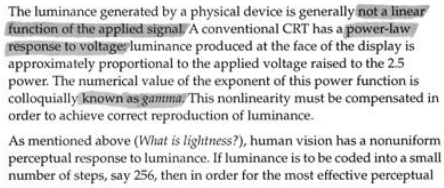
0c_y4.png



The luminance generated by a physical device is generally not a linear function of the applied signal. A conventional CRT has a power-law response to voltage: luminance produced at the face of the display is approximately proportional to the applied voltage raised to the 2.5 power. The numerical value of the exponent of this power function is colloquially known as *gamma*. This nonlinearity must be compensated in order to achieve correct reproduction of luminance.

As mentioned above (*What is lightness?*), human vision has a nonuniform perceptual response to luminance. If luminance is to be coded into a small number of steps, say 256, then in order for the most effective perceptual

0c_y5.png



The luminance generated by a physical device is generally not a linear function of the applied signal. A conventional CRT has a power-law response to voltage: luminance produced at the face of the display is approximately proportional to the applied voltage raised to the 2.5 power. The numerical value of the exponent of this power function is colloquially known as *gamma*. This nonlinearity must be compensated in order to achieve correct reproduction of luminance.

As mentioned above (*What is lightness?*), human vision has a nonuniform perceptual response to luminance. If luminance is to be coded into a small number of steps, say 256, then in order for the most effective perceptual

0c_y6.png

(5) Discussions

The results in 0a is out of my expectation since the weight of green seems stronger than that of red. All the 2D surfaces are similar to Fig 2.

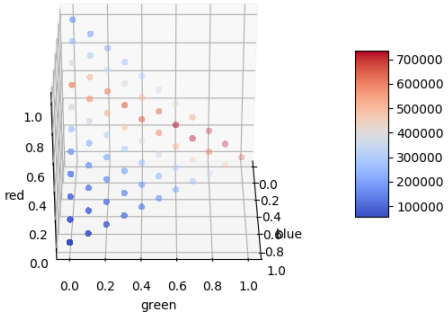


Fig 2. A 2D surface of candidates images for 0a.

For 0b, since the original image is majorly composed of red and green, I think the results are reasonable.

For 0c, 5 candidate images of 0c have tied ballot, which is out of my expectation. A possible reason is that the local minimum is present between the candidates so that all the neighboring candidates are similarly possible to be the answer. For example, (0.3, 0.4, 0.3), (0.3, 0.5, 0.2), and (0.3, 0.3, 0.4) are close to each other.