Computer Vision Homework 1

Luka Eerens
lte@andrew.cmu.edu

16720
CMU

September 12, 2018

Theory Questions
Q 2.1
For Hough Transforms, each image edge votes for the best possible line model.

Consider the equation of a straight line:

$$y = mx + b$$

In the case of this problem, we are given the equation:

$$xcos\theta + ysin\theta - \rho = 0$$

Assume that Hough space transforms things in the image space into a sinusoid of a certain amplitude and phase shift. This is mathematically described by the following equation:

$$\rho = Acos(\theta - \omega)$$

Where A is of course the amplitude, and ω is the phase shift.

Now, $cos(\theta - \omega)$ can be re-written as:

$$cos(\theta)\cos(\omega) + \sin(\theta)\sin(\omega)$$

$$\therefore Acos(\theta - \omega) = A[cos(\theta)\cos(\omega) + \sin(\theta)\sin(\omega)]$$

$$= Acos(\theta)\cos(\omega) + A\sin(\theta)\sin(\omega)$$

We can assume that the x and y components of the sinusoid can be expressed as:

$$x = A\cos(\omega)$$
$$\mathrm{y} = \mathrm{A}\sin(\omega)$$

Which when looking back at the equation given for this problem gives:

$$A\cos(\omega)cos\theta + \mathrm{A}\sin(\omega)sin\theta - \rho = 0$$

Which means:

$$xcos\theta + ysin\theta = \rho$$

Now as far as founding out what the 2 parameter values of Amplitude (A) and phase shift ω are, amplitude is found through Pythagoras:

$$A = \sqrt{x^2 + y^2}$$

Therefore knowing Amplitude, the following equation:

$$\rho = Acos(\theta - \omega)$$

can be expressed as:

$$\rho = \sqrt{x^2 + y^2}cos(\theta - \omega)$$

And to find the phase shift ω, we just need to use the following formula:

$$\omega = \tan^{-1}\left(\frac{y}{x}\right)$$

Therefore we end up getting:

$$\rho = \sqrt{x^2 + y^2}\cos\left[\theta - tan^{-1}\left(\frac{y}{x}\right)\right]$$

Q 2.2

There is practical reason for parameterizing the line in terms of ρ and θ instead of intercepts. Let's have a look at it.

As mentioned in the previous problem, the equation for a straight line is obviously

$$y = mx + c$$

Now from the previous problem:

$$\rho = \sqrt{x^2 + y^2}\cos\left[\theta - tan^{-1}\left(\frac{y}{x}\right)\right]$$

Can be expressed as:

$$\frac{\rho}{\sqrt{x^2 + y^2}} = \cos\left[\theta - tan^{-1}\left(\frac{y}{x}\right)\right]$$

You are then able to reorganize this equation such as:

$$\frac{\rho}{\sin(\theta)} = -xcot(\theta)$$

The big problem with expressing in terms of M and X for a linear equation is that you in the cases where the lines are vertical you are get unbounded values. In this case, having vertical lines means that $\theta = 90$ degrees which when you plug it into the equation above, you do have bounded values.

Bounded values are less computationally taxing to work with and this is why we choose to parameterize the line instead of expressing it through a linear equation.

Now looking back to a linear equation: $y = mx + c$

We see that we get:

$$y = \ -xcot(\theta) + \frac{\rho}{\sin(\theta)}$$

Where m = $xcot(\theta)$ and c = $\frac{\rho}{\sin(\theta)}$

Q2.3

From the 2 previous problems, we obtained the following equation for ρ at some point.

$$\rho = \sqrt{x^2 + y^2} \cos\left[\theta - tan^{-1}\left(\frac{y}{x}\right)\right]$$

Therefore, the maximum value that it can take on is when the cos[…] = 1.
Therefore, the maximum value is:

$$\rho = \sqrt{x^2 + y^2} * 1$$

And since the width (x) of the image is W and the height (y) of the image is H:

$$x = W$$
$$y = H$$

Therefore, subbing in the values yields:

$$\rho = \sqrt{W^2 + H^2}$$

So the largest possible ρ is $\sqrt{W^2 + H^2}$, now to find the smallest, so that we can capture the full range of possible ρ values, we need to go back to the equation at the beginning of this problem.

To get the smallest possible value, we would need cos[…] = -1 . Therefore following the same logic as before, we get a minimum value of ρ

$$\rho = -\sqrt{W^2 + H^2}$$

Now it could be just left at that where:

$$-\sqrt{W^2 + H^2} \leq \rho \leq \sqrt{W^2 + H^2}$$

However, as disclosed in one of the computer vision lectures, there are 2 ways to write the same line. We thus do not need both a negative value and a positive value of ρ to capture the full range of ρ. We could go with one or the other.

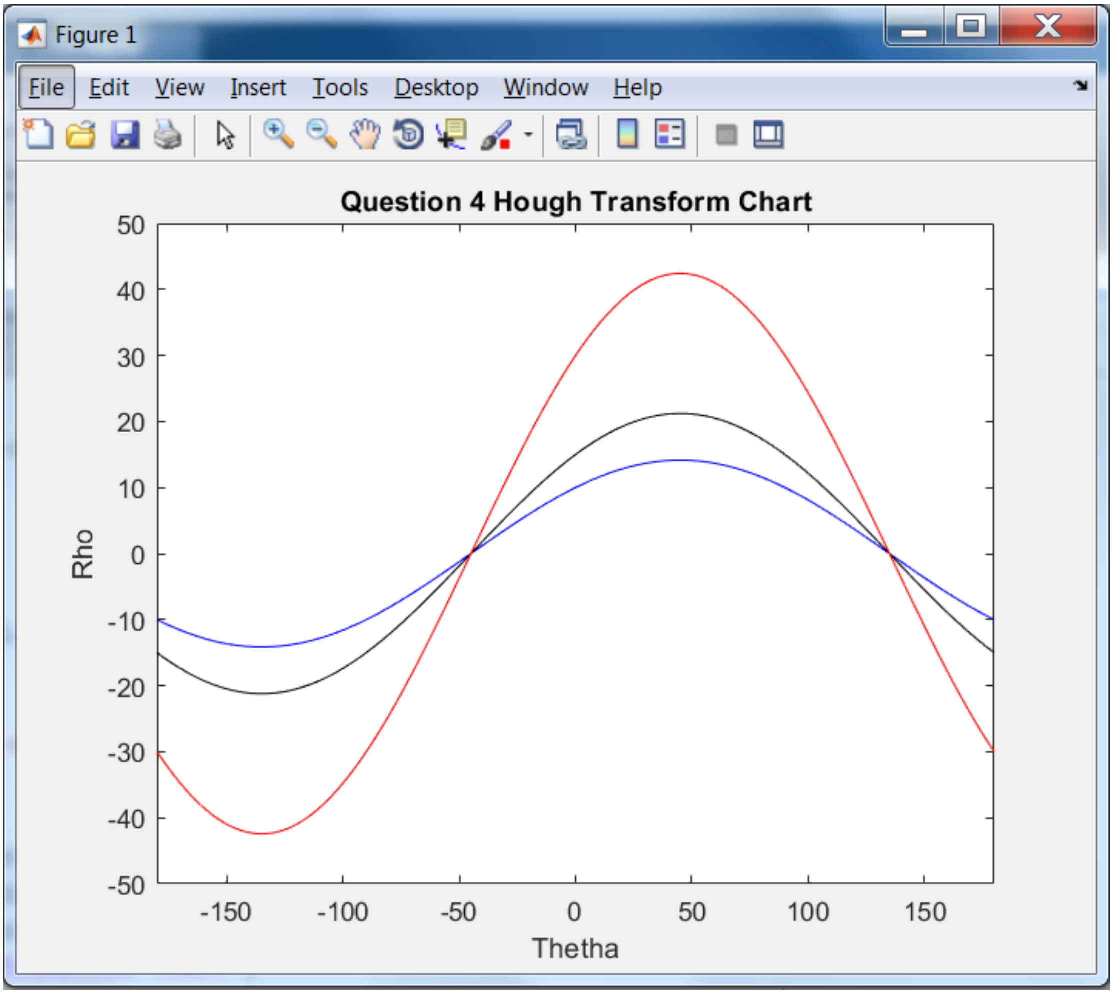
We also do not need to capture the full range of angles (0-360^O) and can constrain it to 0-180^O, due to the similarity in values above and below 180^O.

<u>Q2.4</u>

We are still dealing with the same image referred to in the previous theory questions.

Below is a chart that shows how the corresponding sinusoidal waves in Hough Space of the points:

(10,10) = Blue
(15,15) = Black
(30,30) = Red



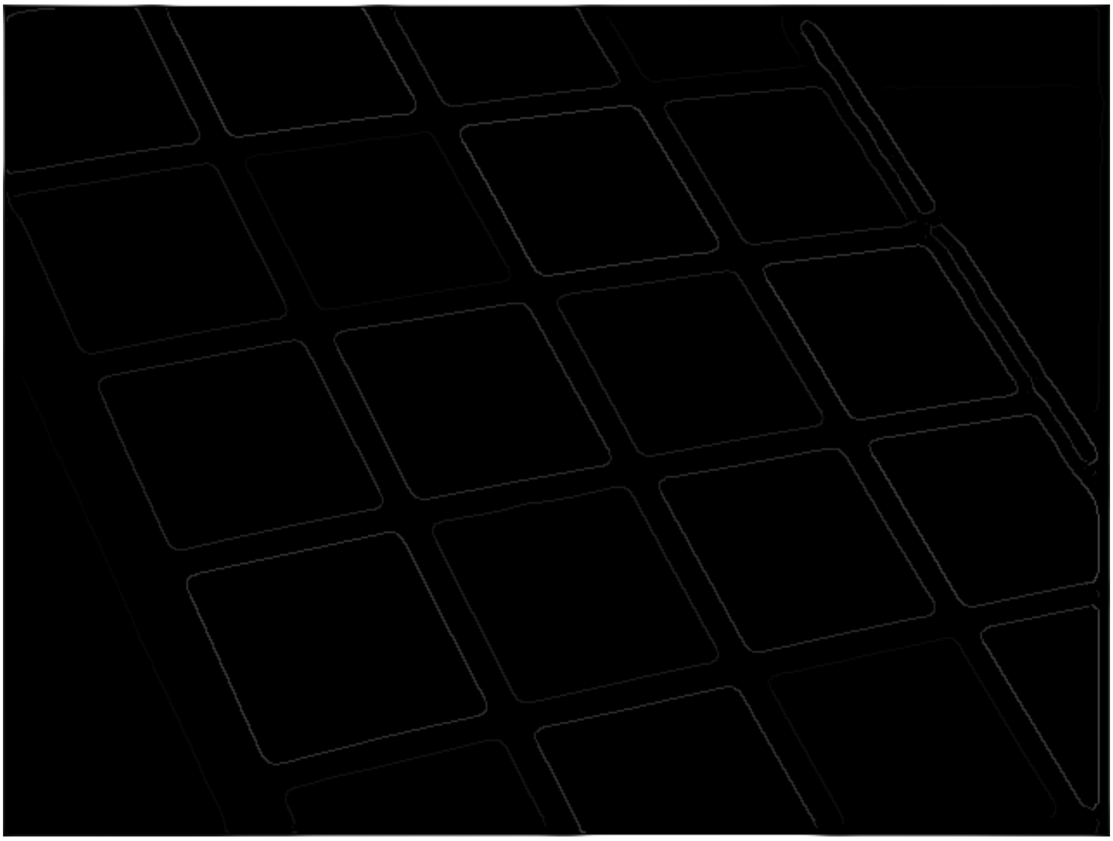
I would guess that the curves meet at a value around -45 degrees for theta and 135 degrees so the expression would be: -45 + 180k where k is an integer. It also looks like they all meet at rho value of 0 for each sight. So Theta = -45 + 180k for any k value iff k is an integer, and Rho = 0.

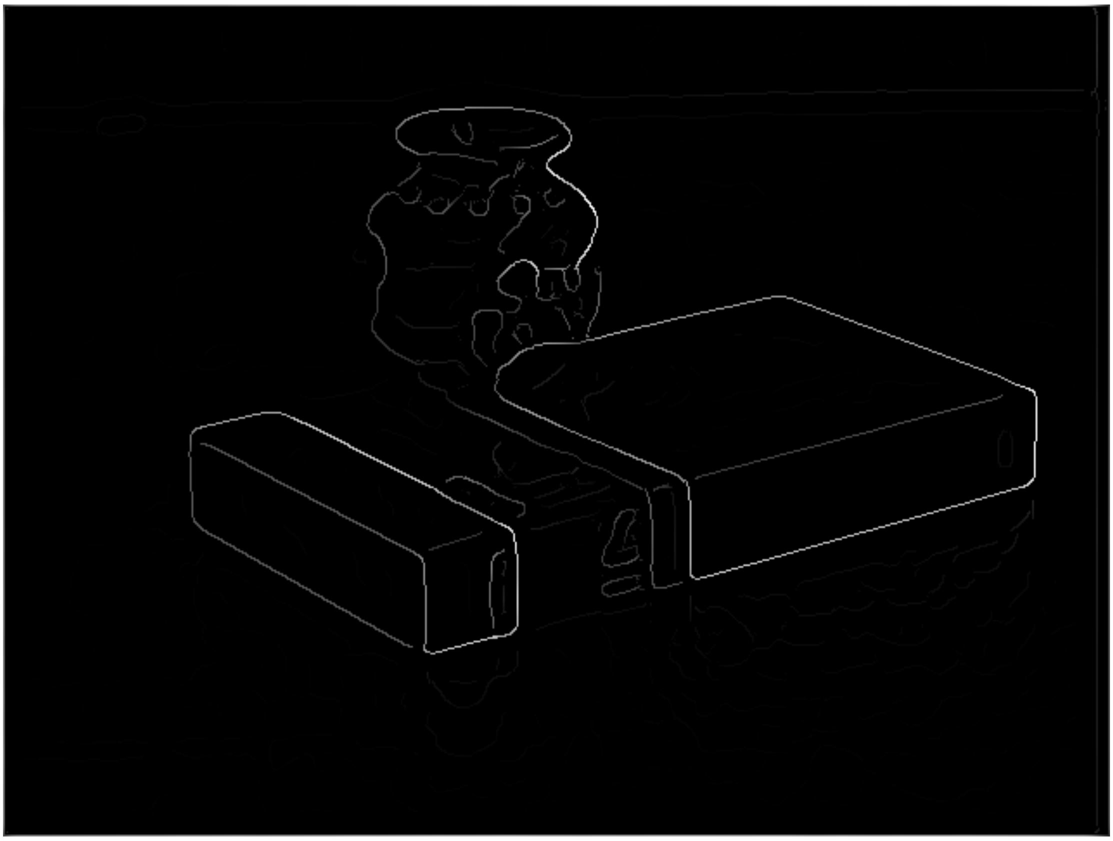
Finding the parameter values m and C when theta and rho are 45 degrees and 0 respectively gives (Where m = $xcot(\theta)$ and c = $\frac{\rho}{\sin(\theta)}$)

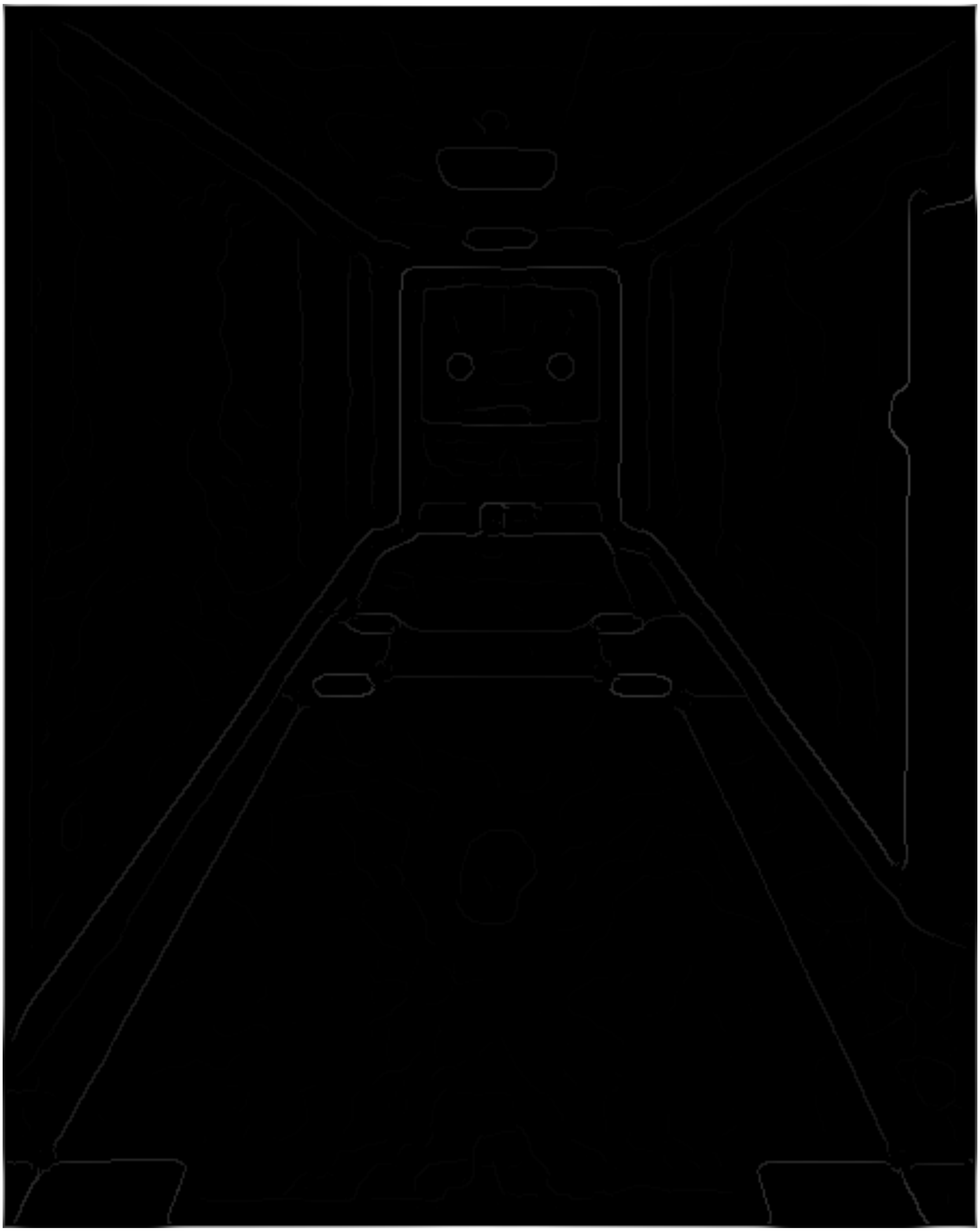
$$b = 0, m = 1$$

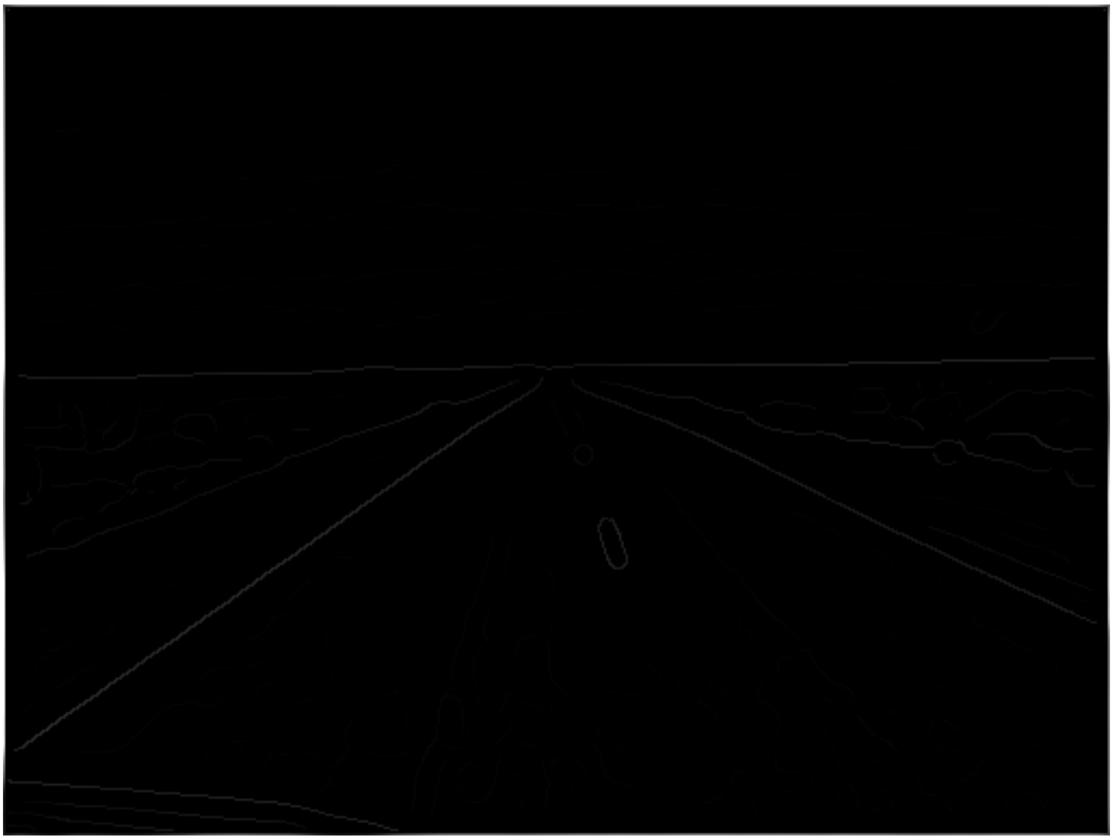
Implementation Questions

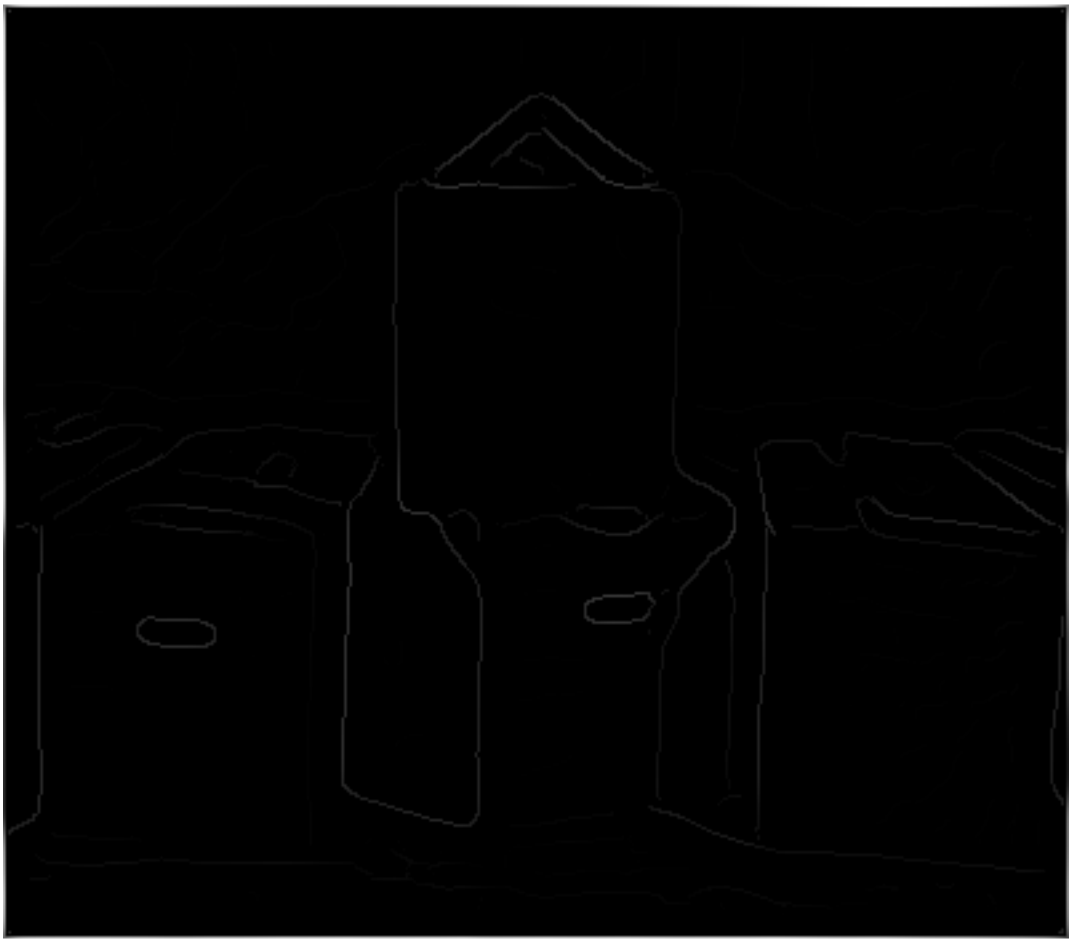
Q3.2 WARNING: ALL OF THE PHOTOS BEFORE ARE FROM THE OPTIMUM PARAM SETTING, THE OUTPUT OF EVERYTHING IS IN THE FOLDER, and THE LINE OUTPUT OF THE DEFAULT VALUES CAN BE SEEN IN THE EXPERIMENT SECTION ALONG WITH THOSE OF THE OPTIMUM PARAM SETUP

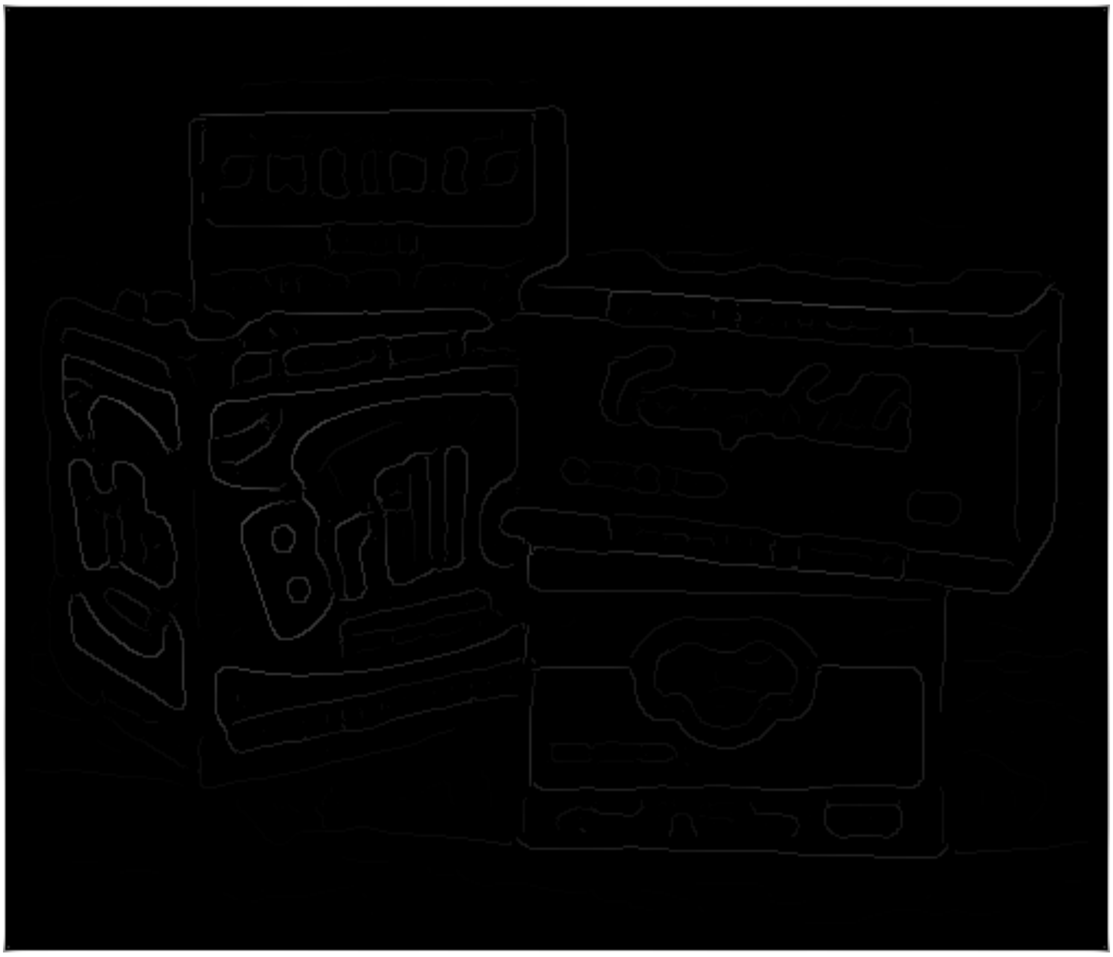


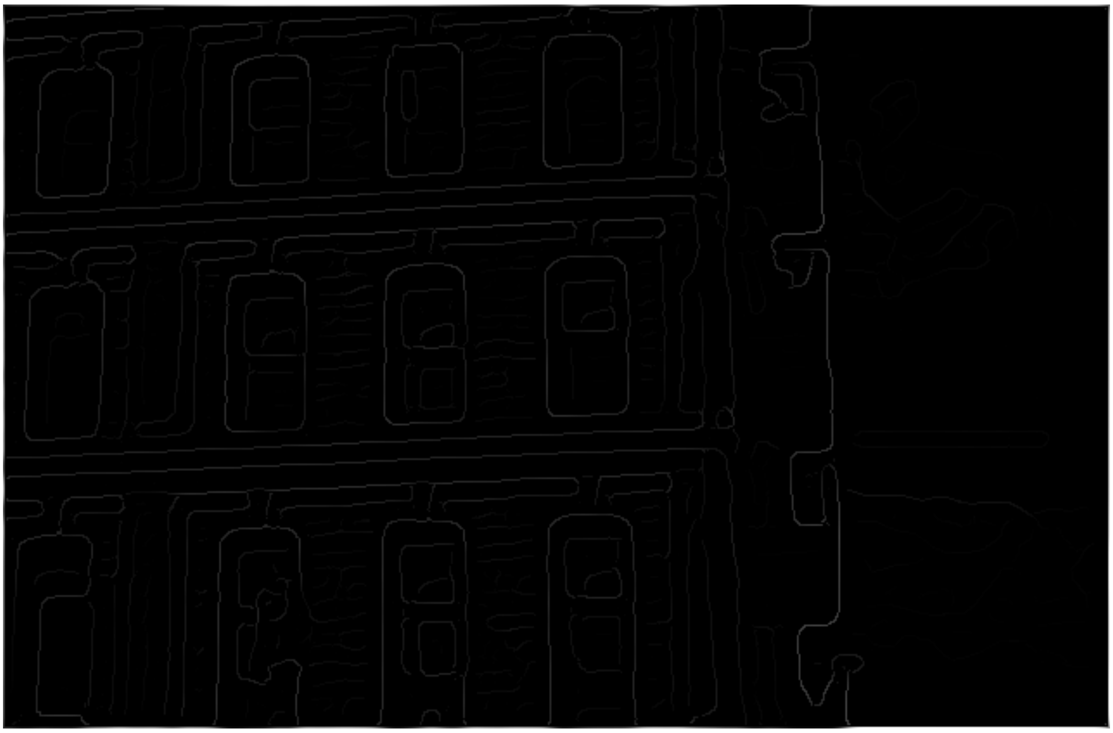




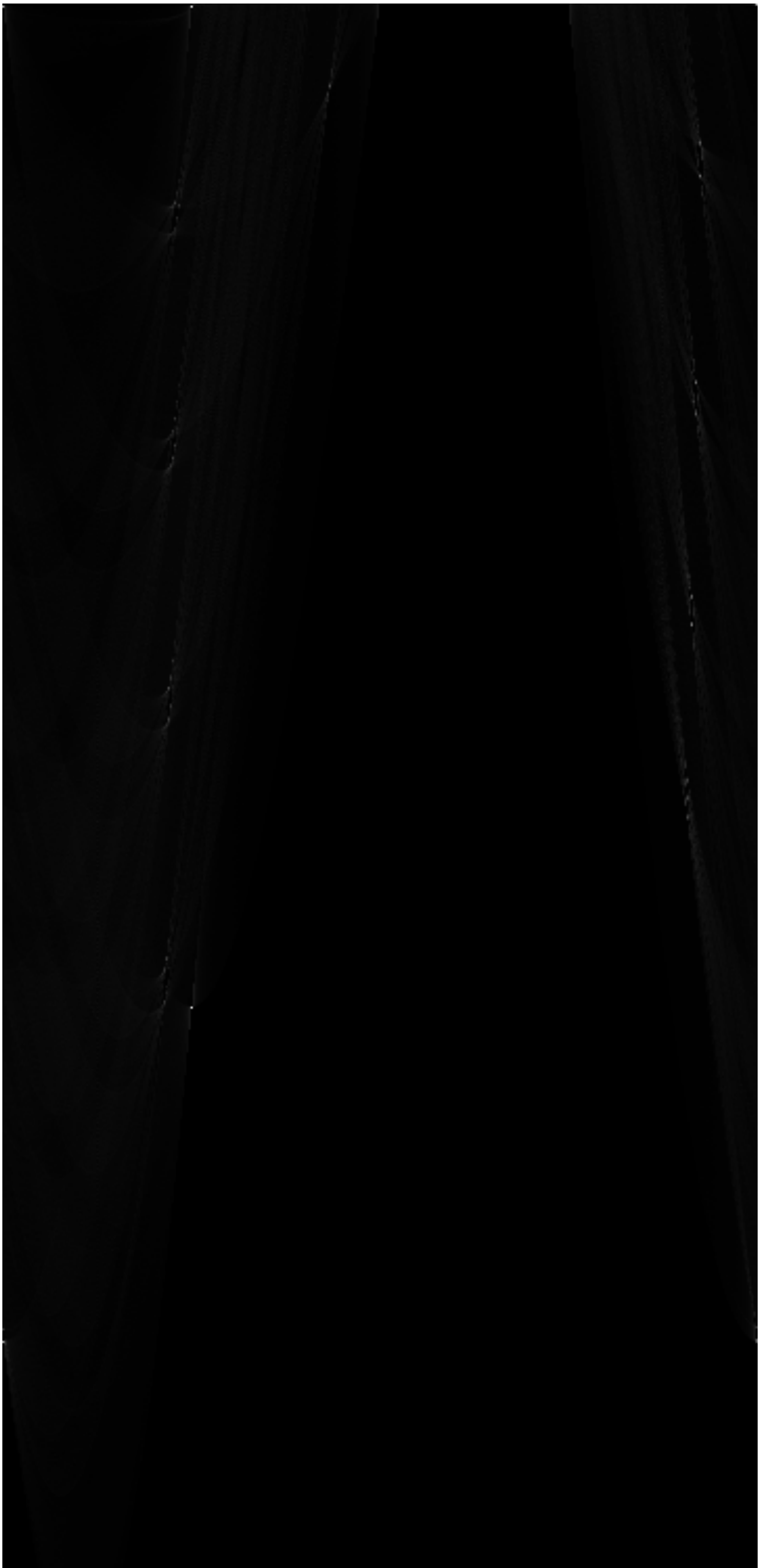


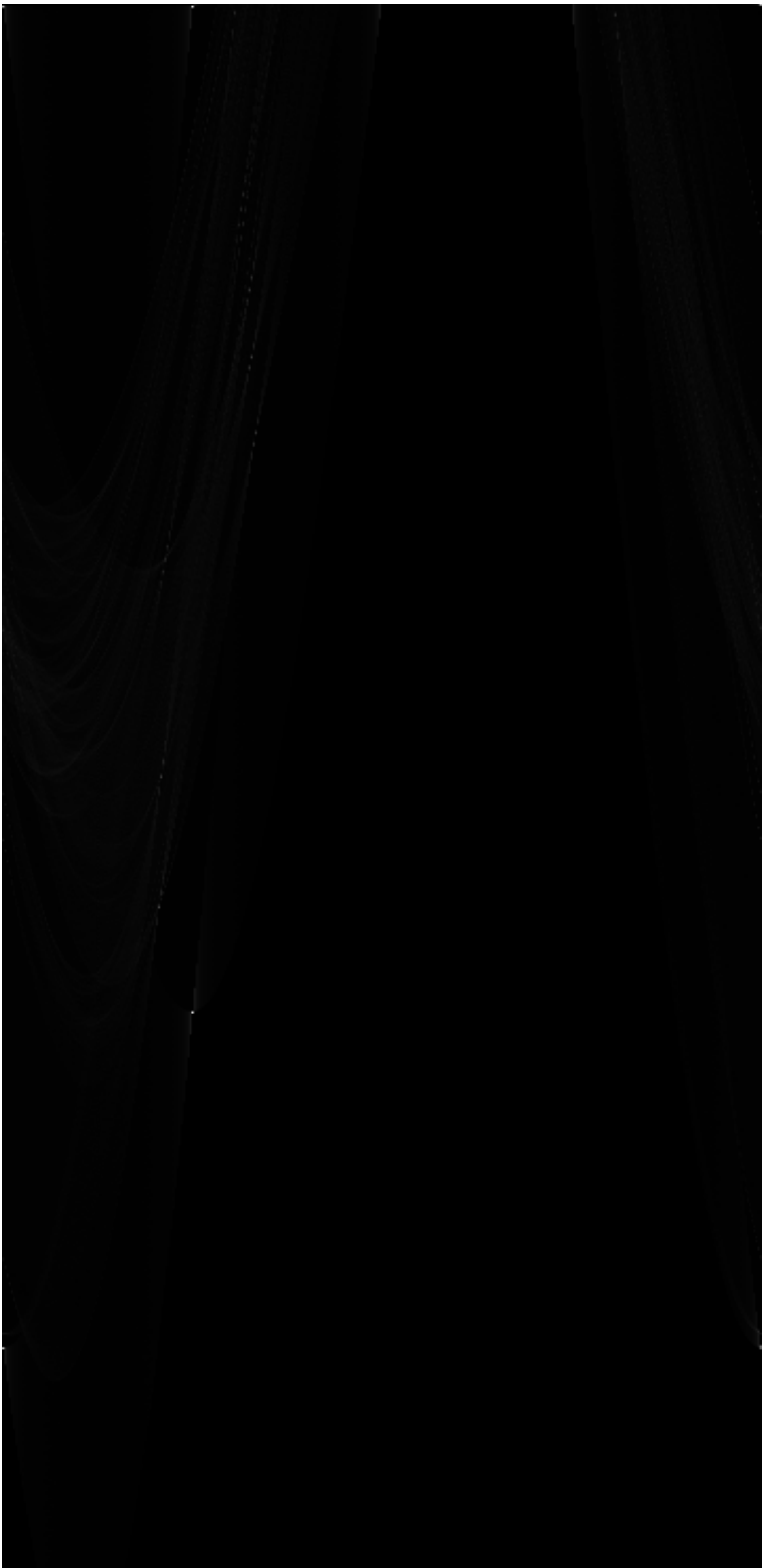


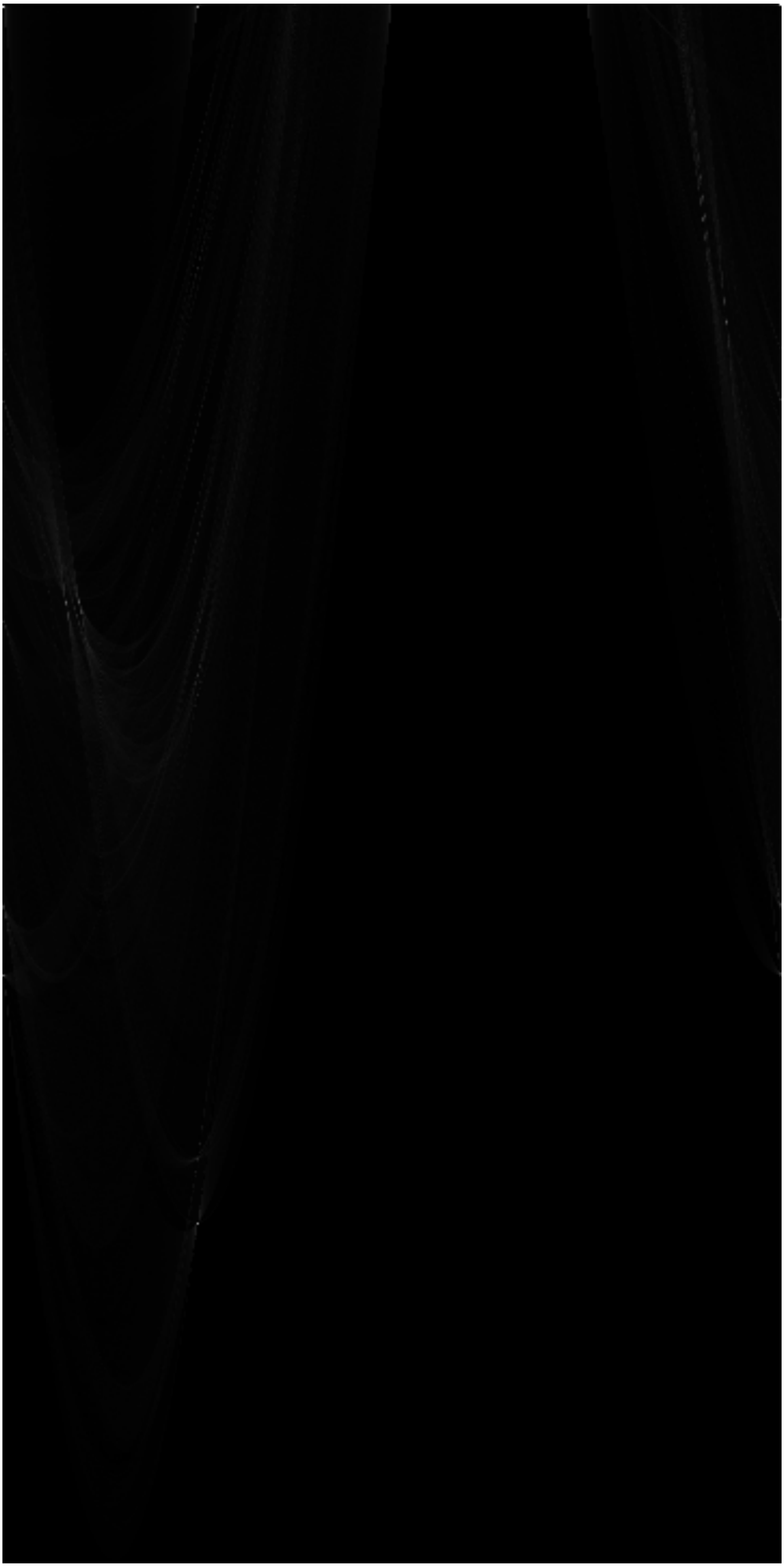


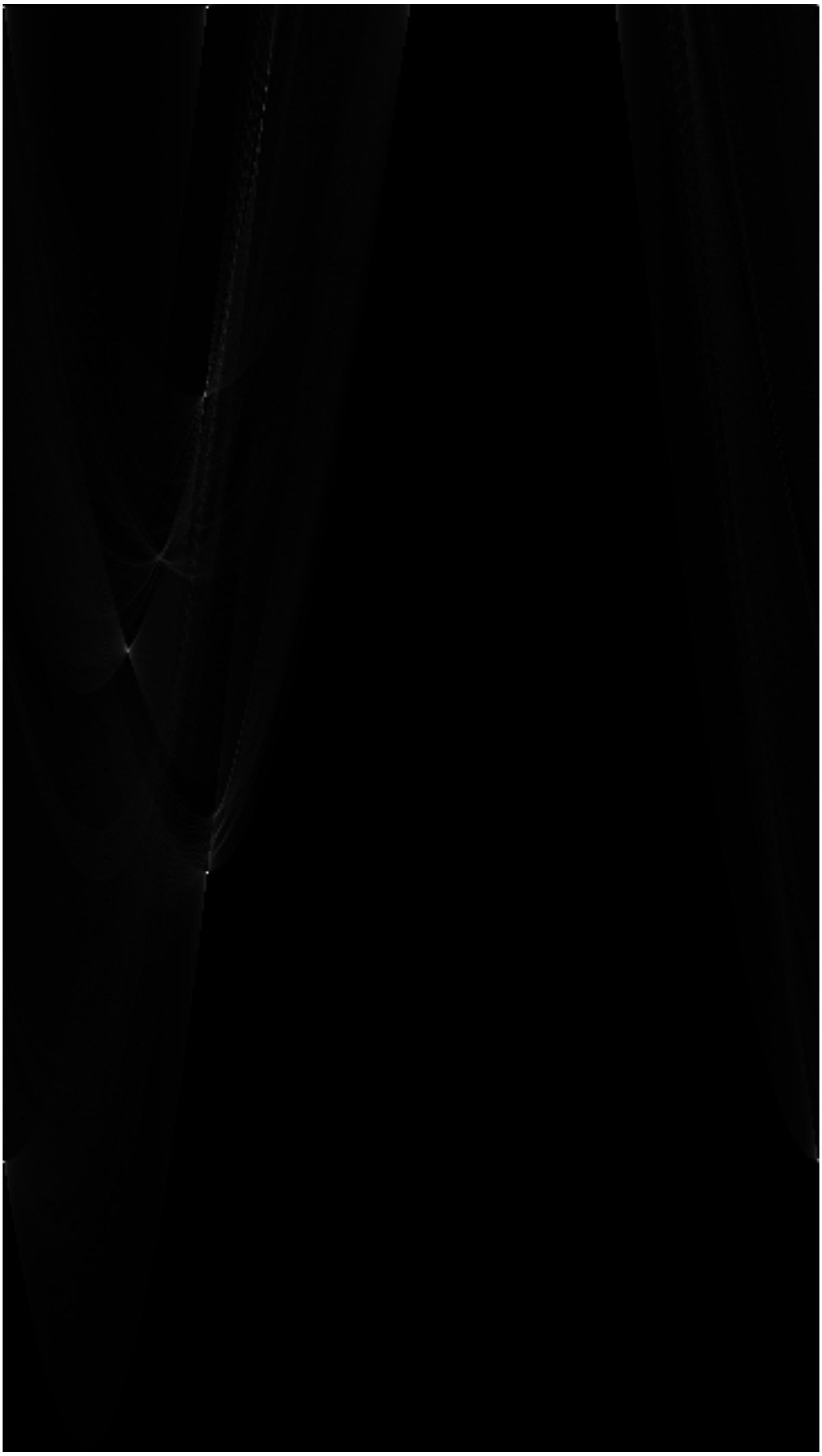


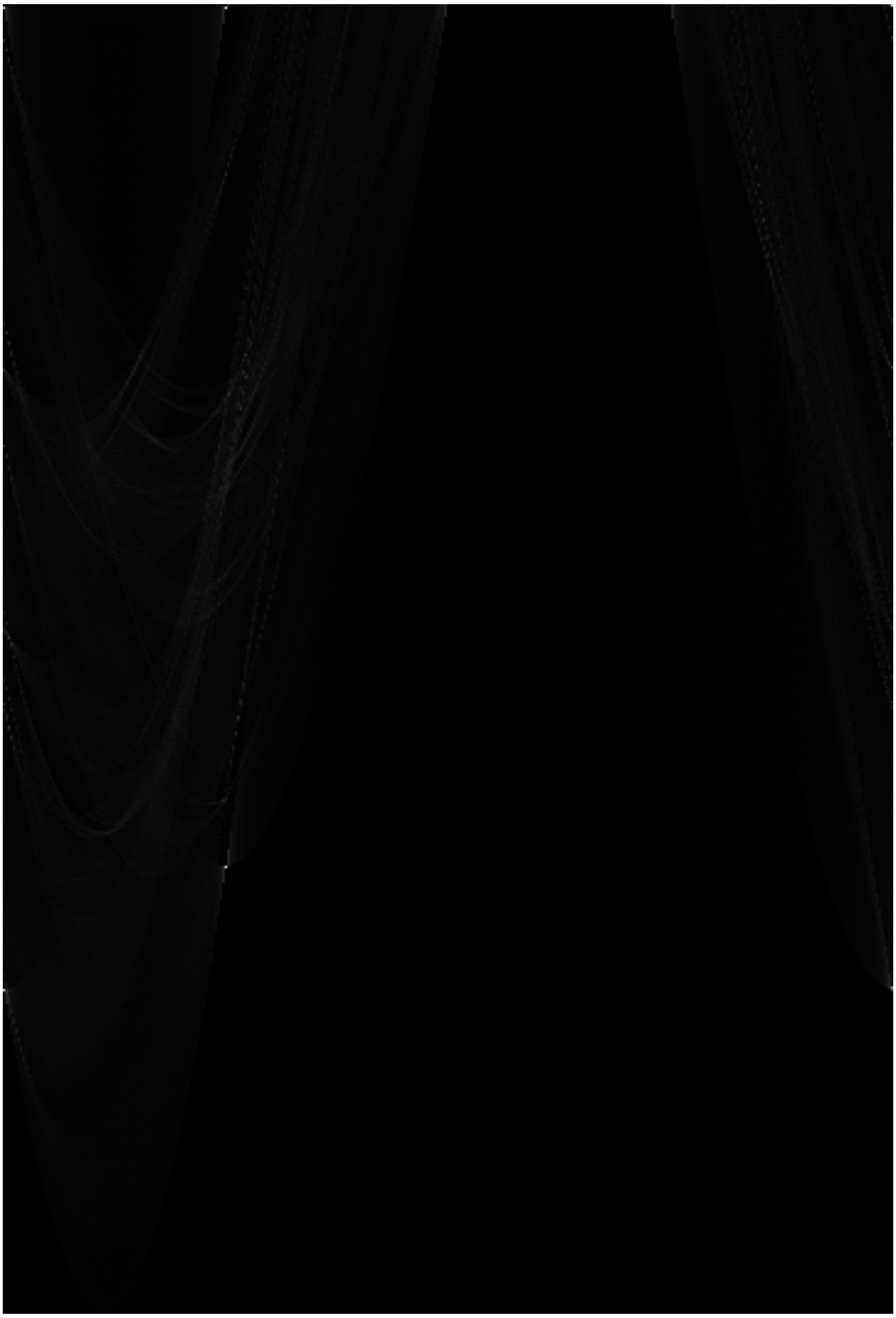
Q3.3

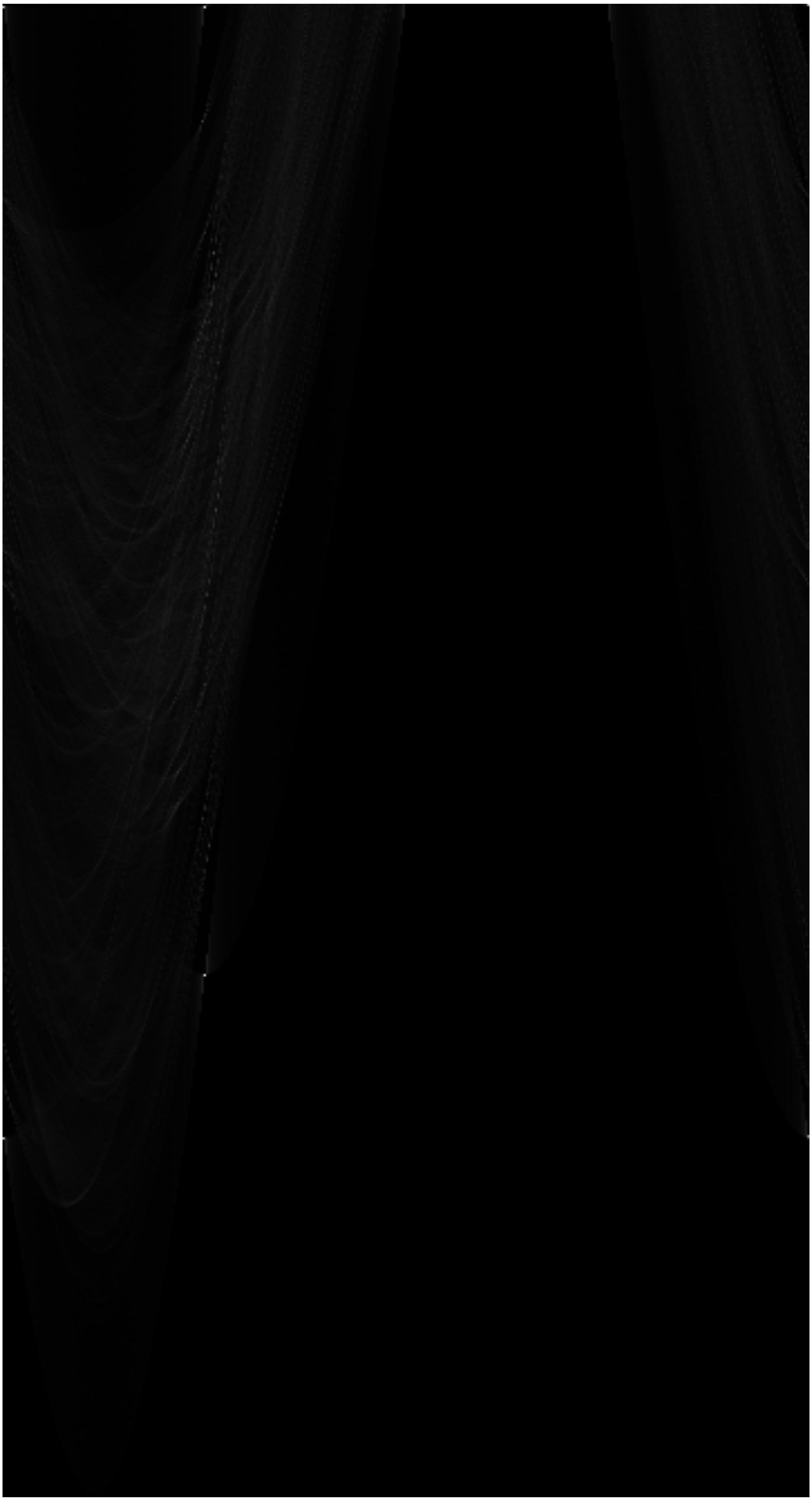


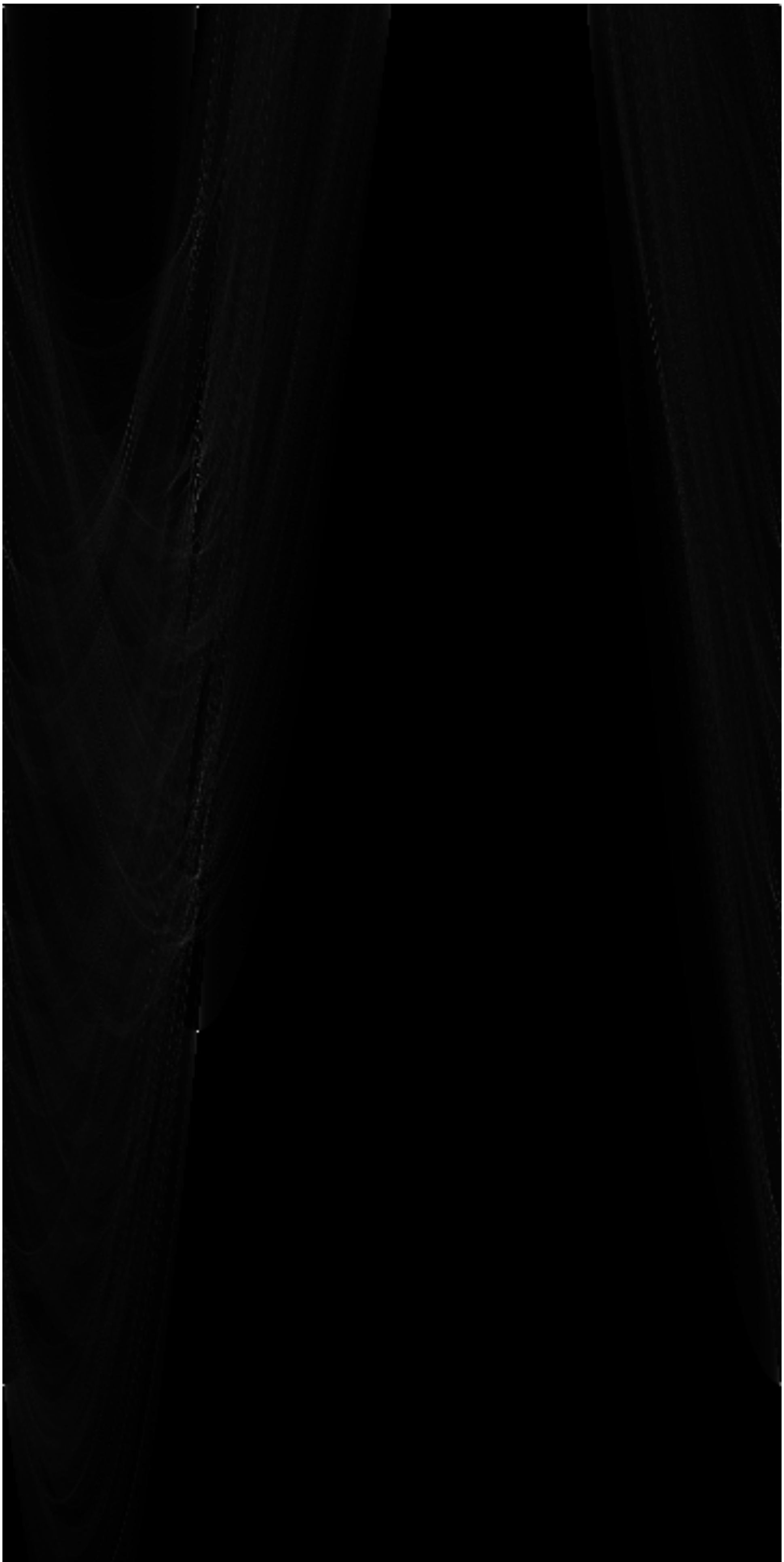


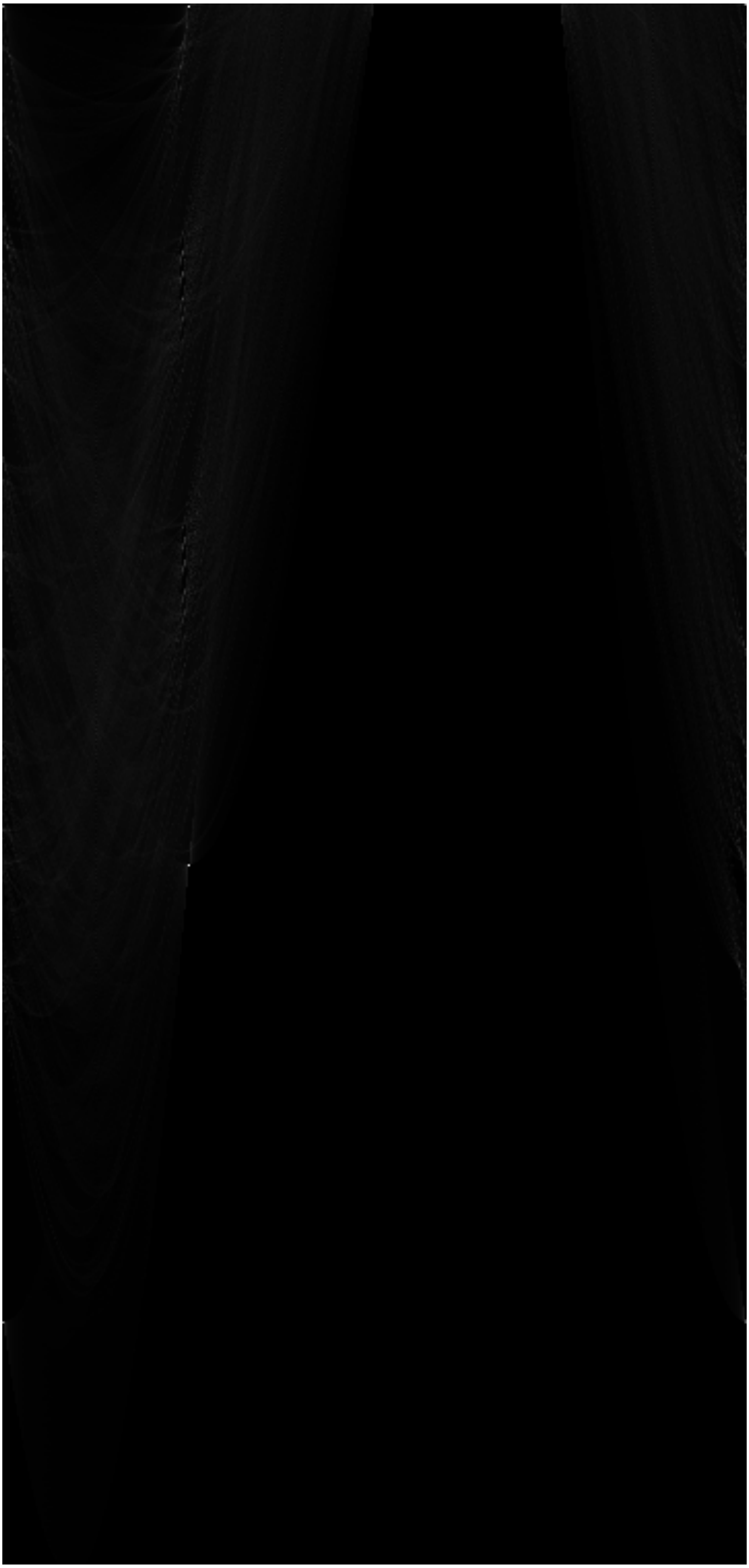


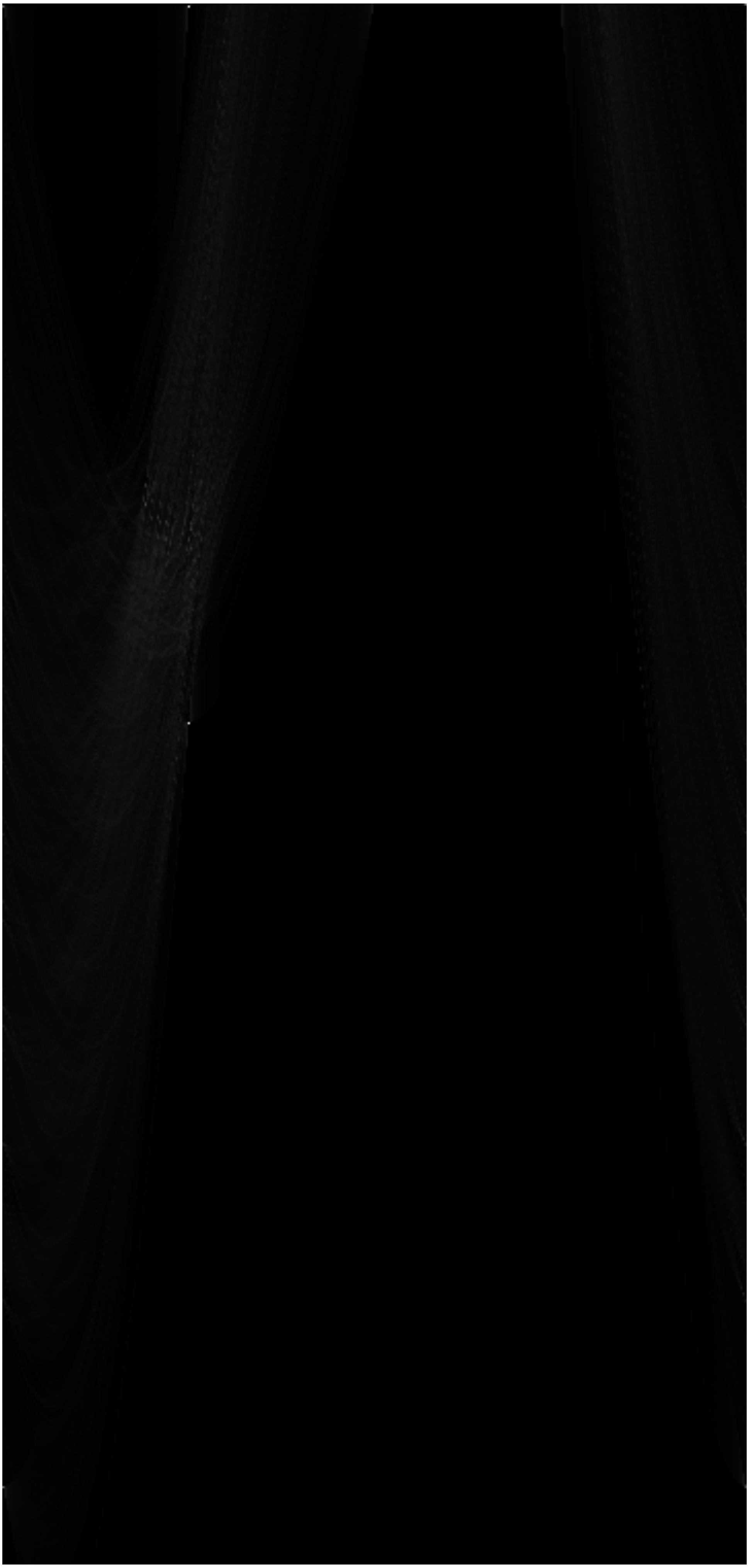




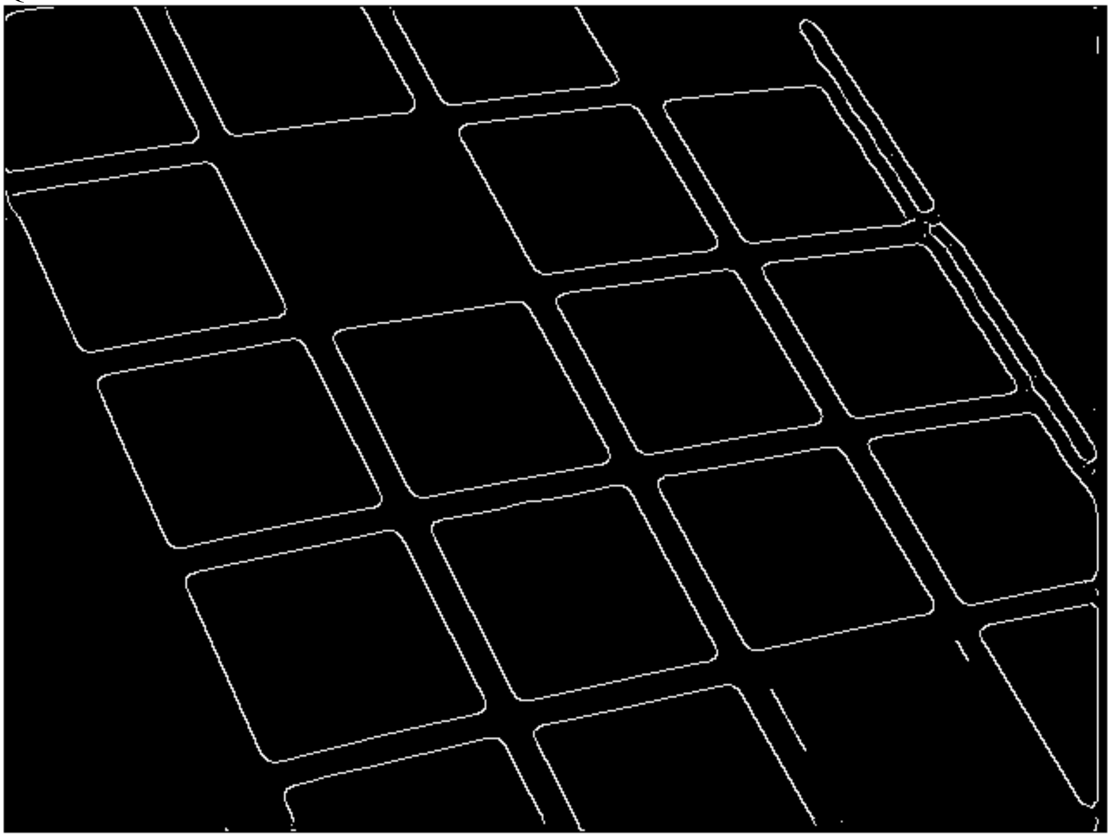


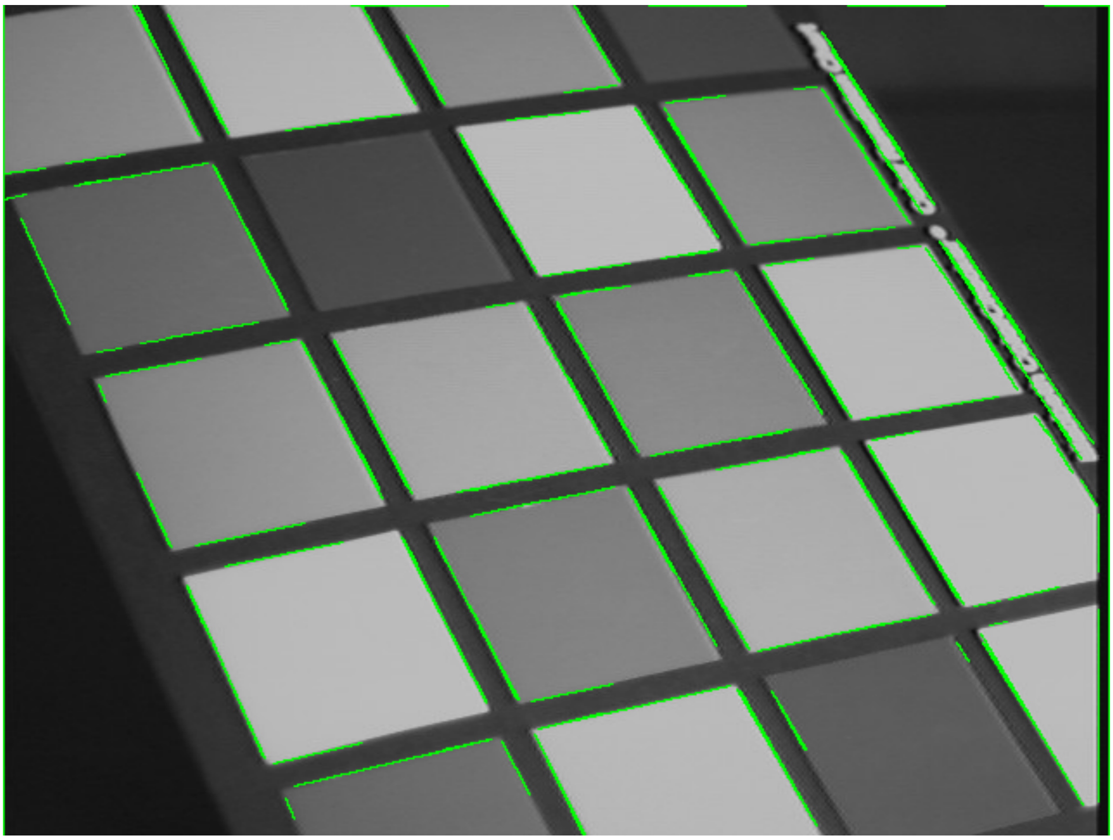


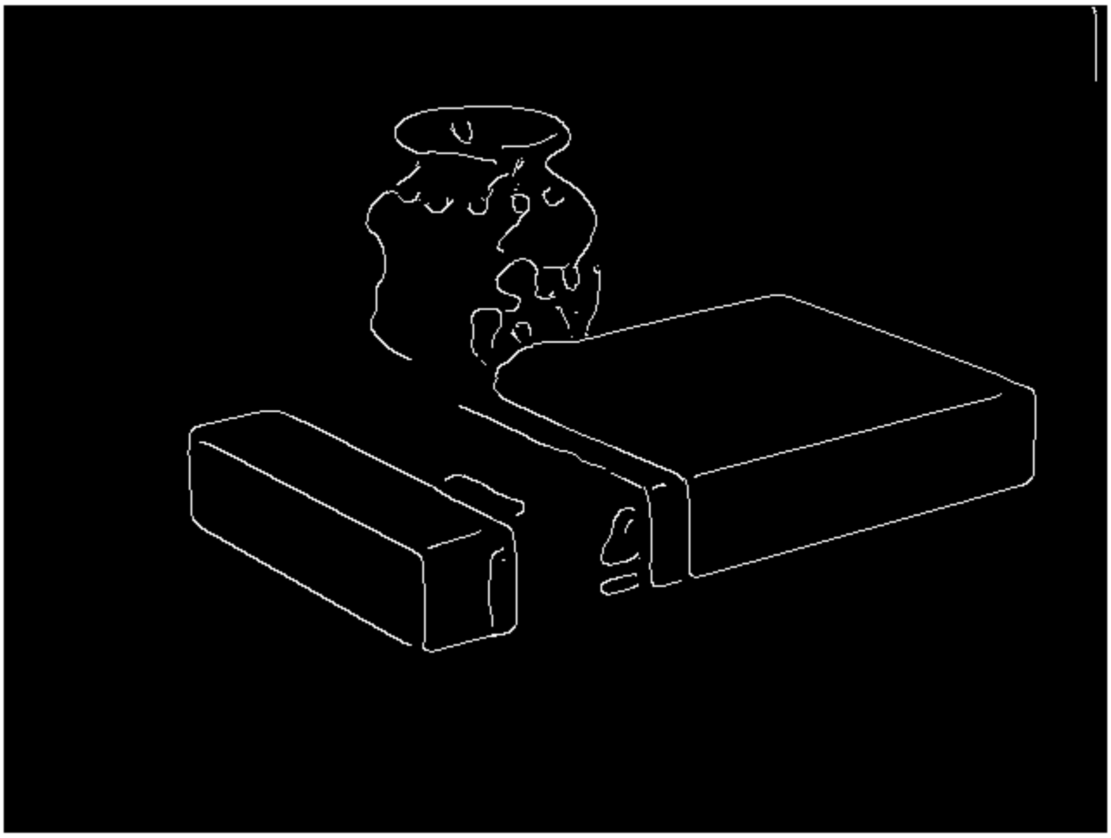


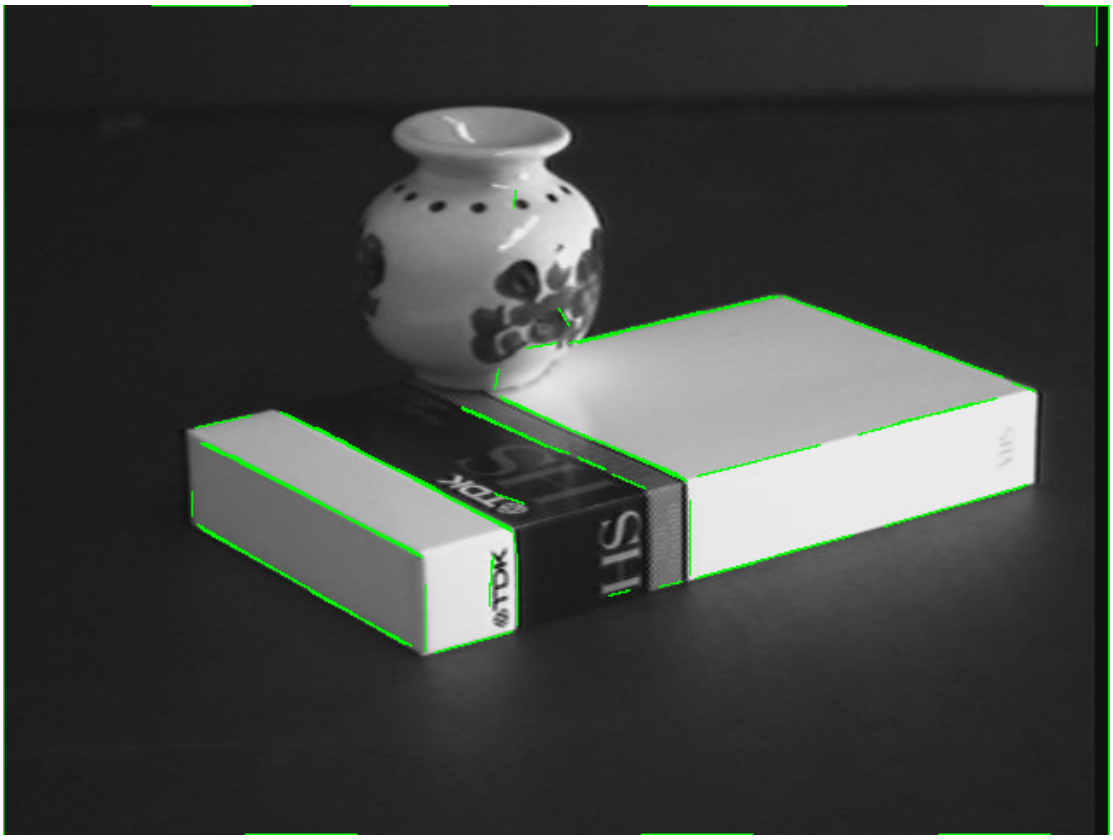


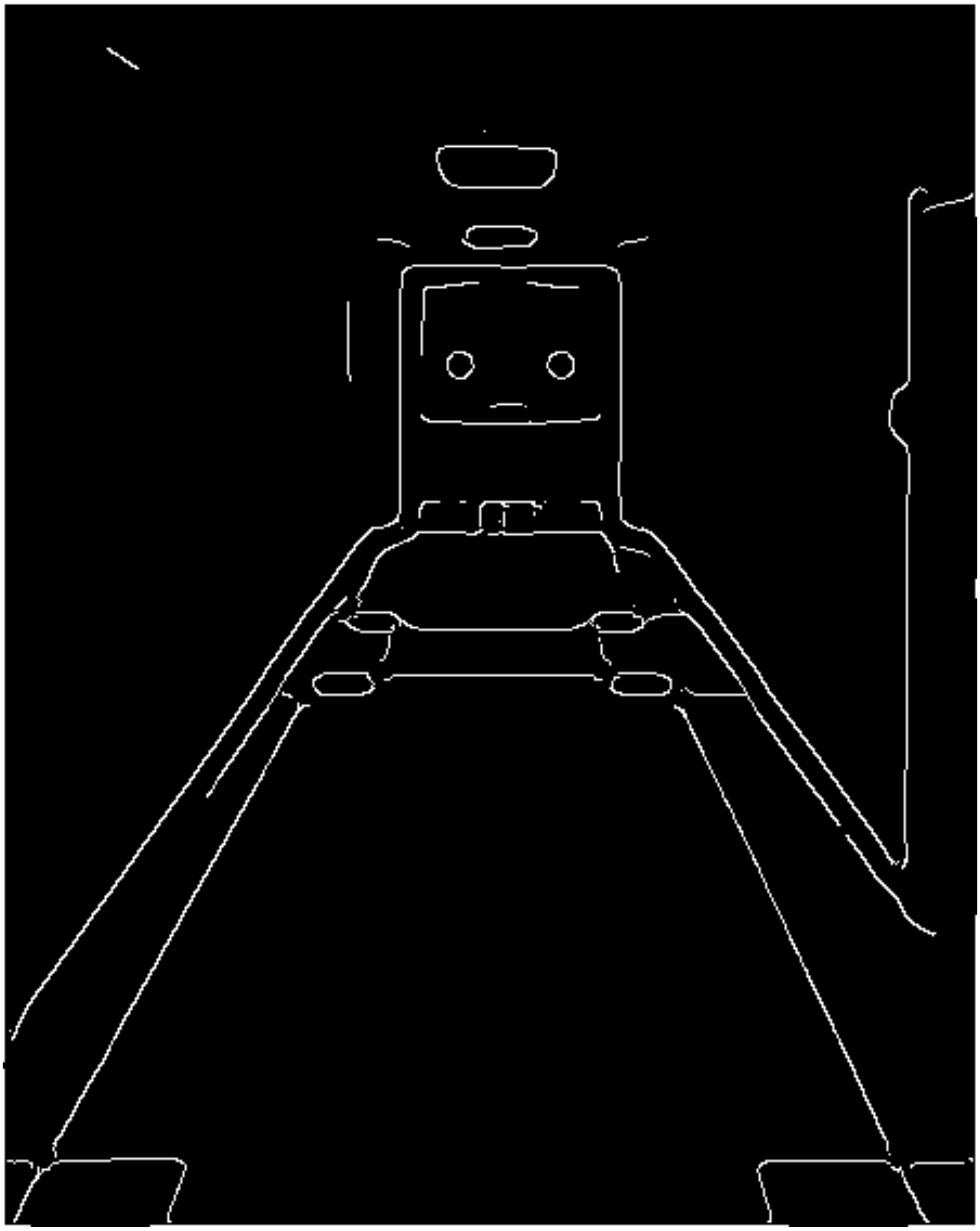
Q3.5

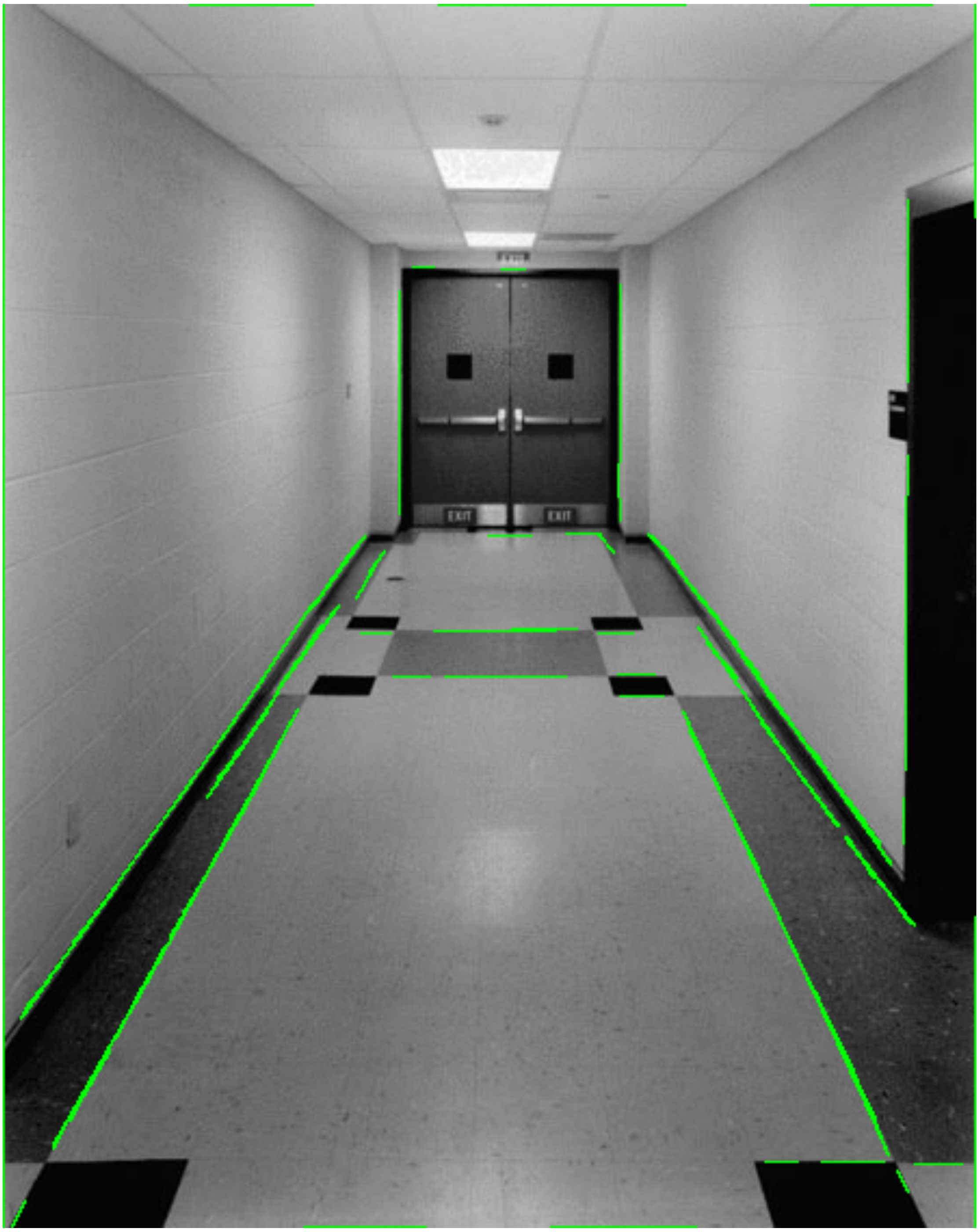


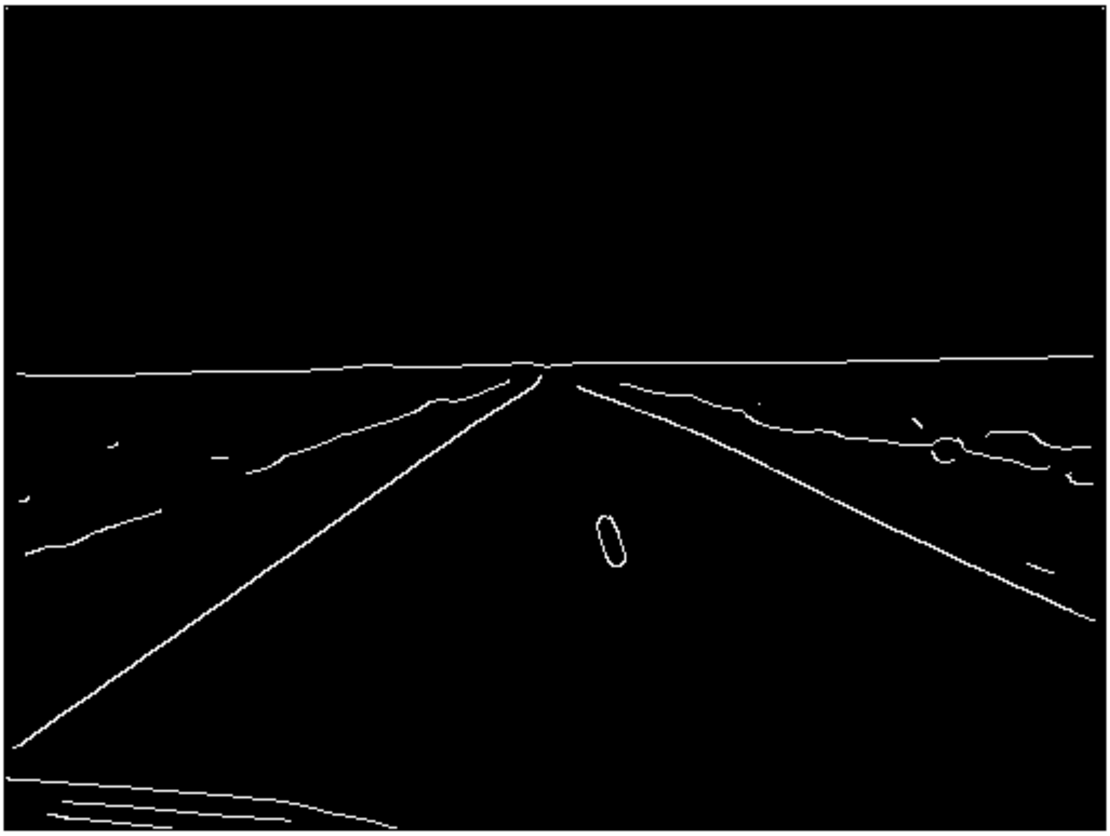


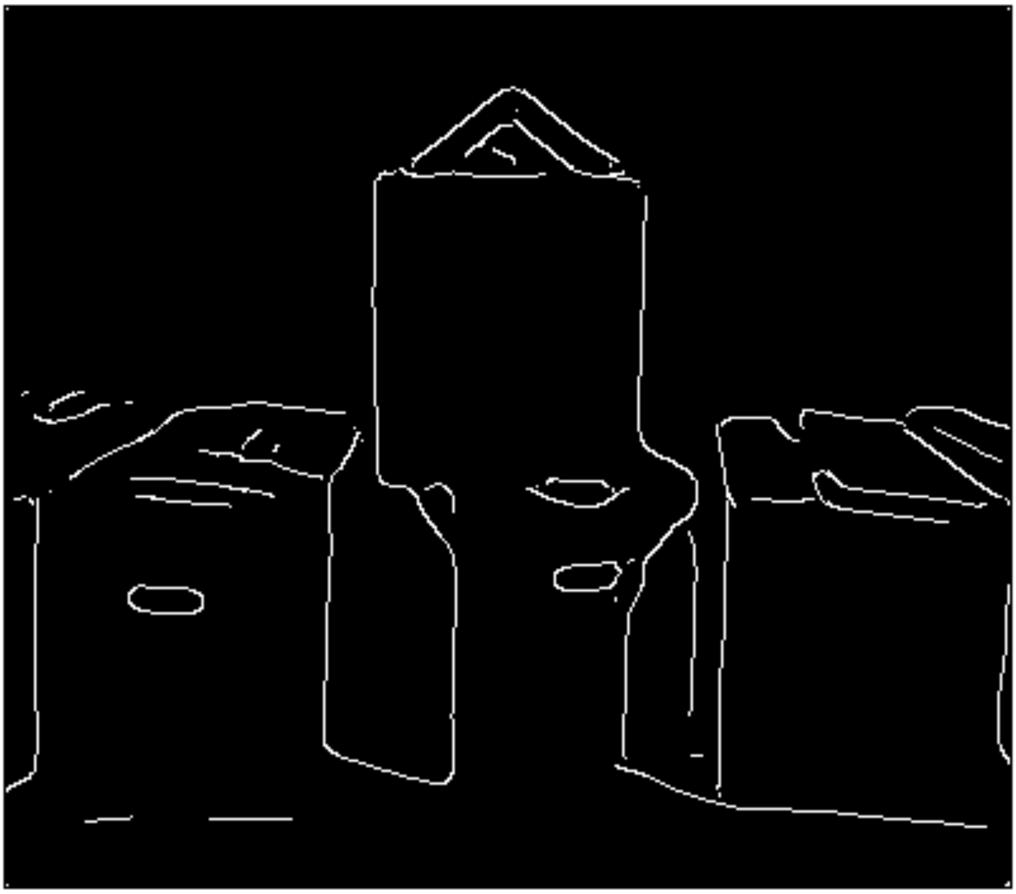


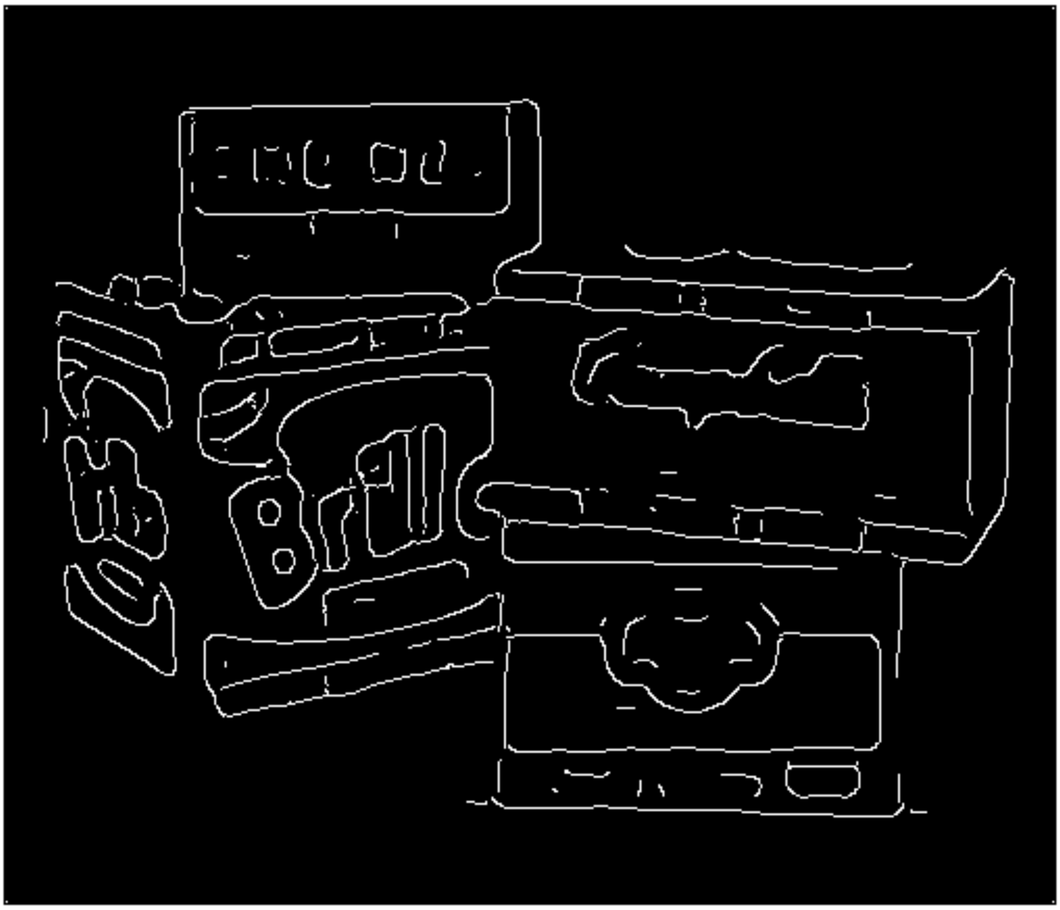

TDK



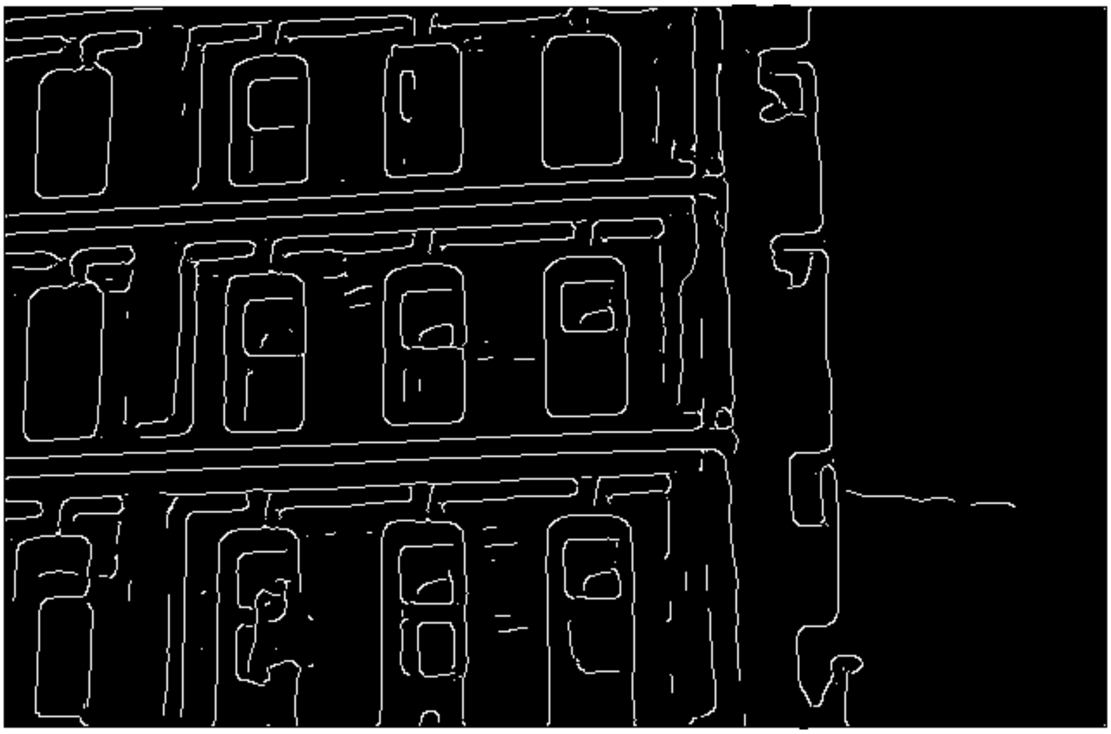









The World's Largest Selling Ketchup
57 HEINZ 57
2 DOZEN
14 OZ. BOTTLES
5124
HEINZ TOMATO KETCHU
24 GIANT SIZE PKGS.
24 GIANT SIZE PKG
New!
Brillo.
Brillo
soap pads
WITH RUST RESISTER
SHINES ALUMINUM FAS
TOMATO JUICE
Campbell's
36-17½ OZ. CANS
362
TOMATO JUICE
Del Monte
QUALITY
24 No. 2½ CANS
FREESTONE
PEACH HALVES
1069




GABIRULES54.DEVIANTART.COM

4. Experiments

The code worked well with the single set of default parameters however there were many difficulties in getting it to work properly. I made some initial mistakes, which led a cascade of spurious, false output for all function outputs thereafter. I found that the hardest part of this assignment was the Hough transform in large part because I didn't really grasp the idea of lines converted into parametric form with rho and theta as the axis. To add, I tried to concretely and visually understand it by experimenting with my code, but the output was incorrect because of my faulty definition of my Sobel filters, which ended up outputting utter nonsense. So I would say that the Hough Transform was my biggest pain point.
Thankfully I was able to get it to work in the end pretty well.

I found some results from the variations of the parameters. Namely, by increasing Sigma (the size of the Kernel receptor field) less noise became present. This was true up to a point because detail and granularity is lost through the convoluted sum of an even larger neighborhood of pixels. So the larger the sigma, the less the number of detected edges, as the blurring effect tempers the intensity gradients. Less than 4, there is less noise removal, and so a larger amount of potential edges that may manifest themselves in the interaction between noise and actual edges in the image. So decreasing Sigma, increases the number of lines detected, these being smaller because they are more attributed to noise

Increasing nLines, with no surprise to anyone increases the number of lines detected in the image, these being of no better quality than the ones already there, in fact they seem to worsen in quality.

Below is the actual study of the parameters:

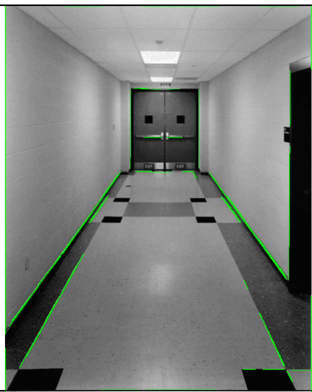
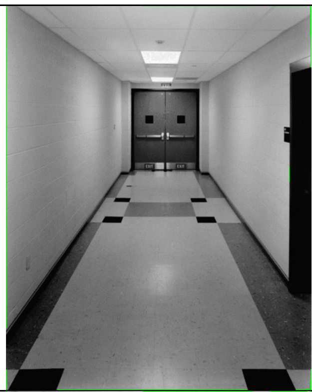
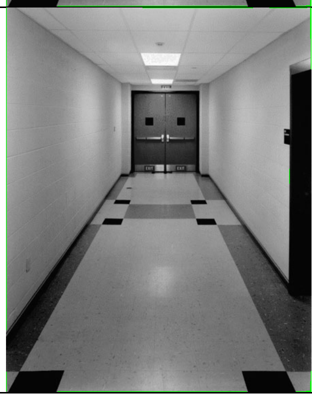
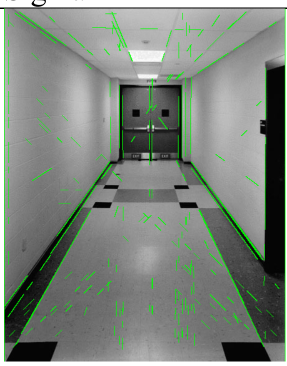
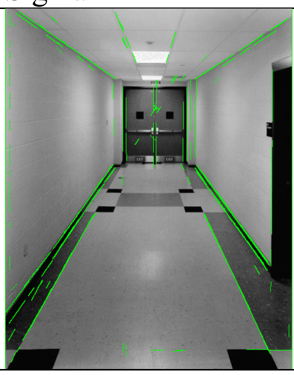
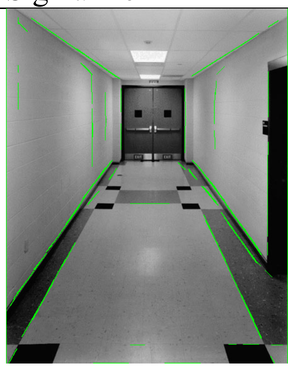
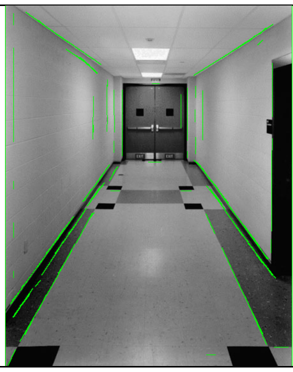
Image 3	Threshold = 0.001	Threshold = 0.03	Threshold = 1
rhoRes = 1			
rhoRes = 2			

Image 3	Sigma = 2	Sigma = 4	Sigma = 6
thetaRes = pi/180			

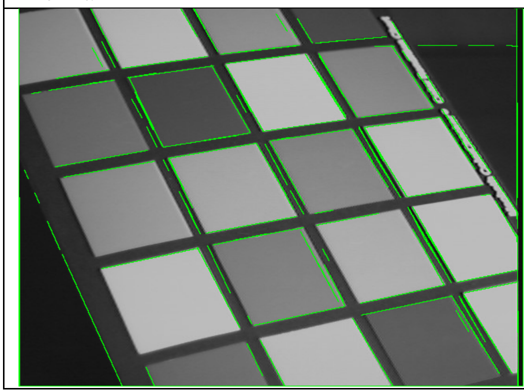
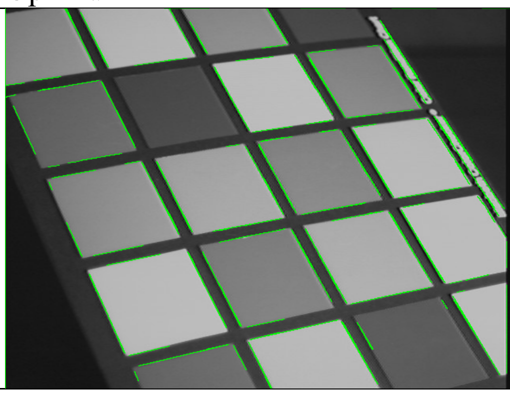
thetaRes = pi/90			

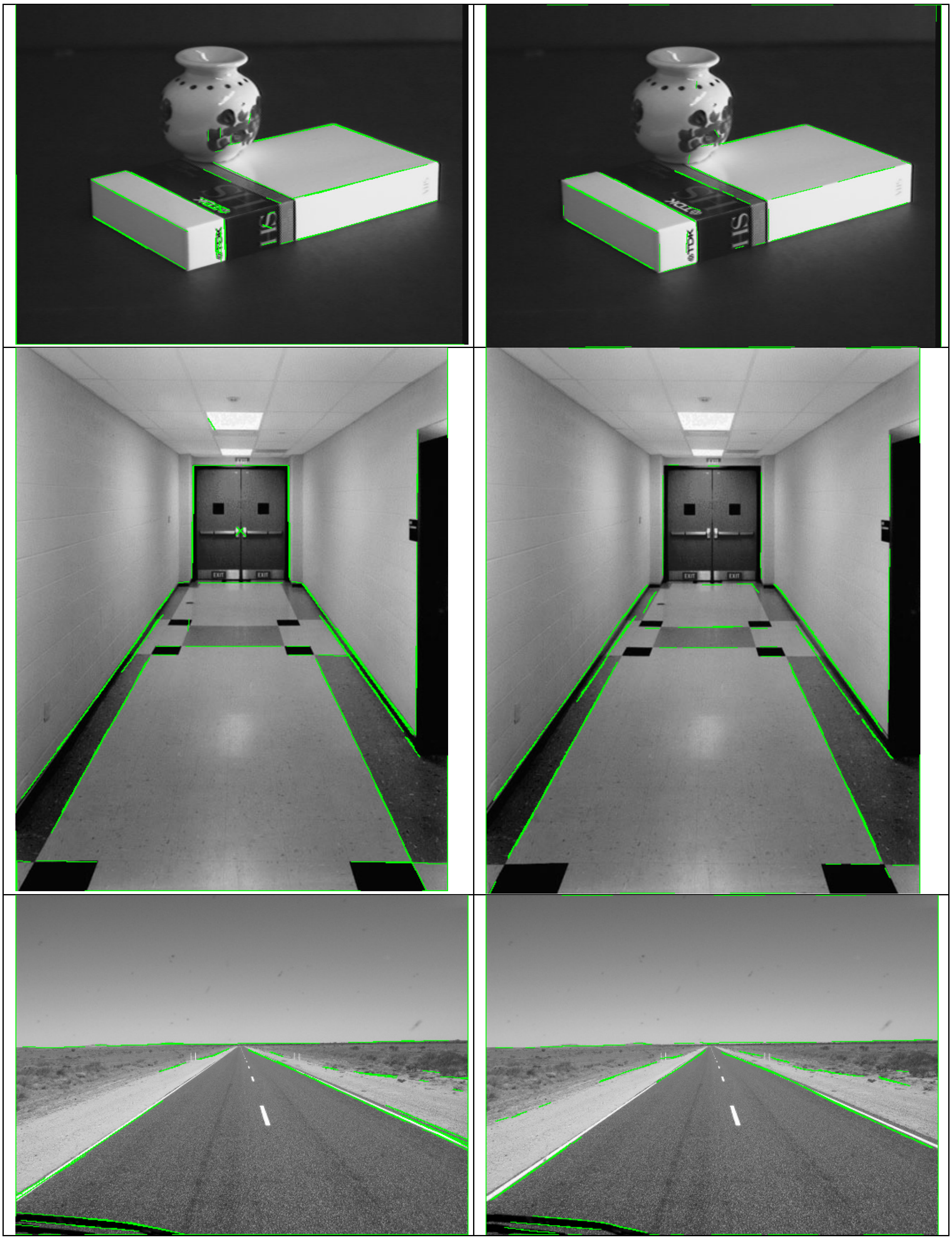
nLines	nLines = 20	nLines = 50	nLines = 100
			

Below is a comparison of the original output with default params, as well as the optimum params:

Here is first of all a side by side comparison of the actual parameters:

	Default	Optimum
Sigma	2	4
Threshold	0.03	0.1
RhoRes	2	1
ThetaRes	Pi/90	Pi/180
nLines	50	100

Default	Optimum
	


TDK


57 HEINZ 57
5124
HEINZ TOMATO KETCHU
TOMATO JUICE
Campbell's
362
24 GIANT SIZE PKG
New!
Brillo
soap pads
WITH RUST RESISTER
SHINES ALUMINUM FAS
Del Monte
QUALITY
FREESTONE
PEACH HALVES
1069


www.dexigner.com
www.dexigner.com