Aryan Dawer

Department of Computer Science

Dartmouth College

January 25, 2024

PSET 1

For this pset, I shared ideas with my fellow classmate Atul Agarwal, however all the work in this document is my own.

1 Theory Questions

Question 1: Sinusoid in Hough Space

Each point (x, y) in the image space can be represented as a sinusoid in the Hough space (ρ, θ) using the line equation $\rho = x\cos\theta + y\sin\theta$. The amplitude of the sinusoid (the distance from the origin to the point) is given by $\sqrt{x^2 + y^2}$, and the phase ϕ is given by $\arctan\left(\frac{y}{x}\right)$.

The line equation in polar coordinates is given by:

$$\rho = x\cos\theta + y\sin\theta \tag{1-1}$$

It relates a point (x, y) in the image to all possible lines that pass through it, with different values of (ρ, θ). As you vary θ, the resulting ρ values trace a sinusoidal curve in the Hough space, where the horizontal axis represents θ and the vertical axis represents ρ. The reason it's sinusoidal is due to the periodic nature of the sine and cosine functions. For a given point (x, y), as you change the angle θ, the values of $\cos\theta$ and $\sin\theta$ vary in a sinusoidal pattern, which means that ρ also changes sinusoidally with θ.

For a given point (x, y), we can express x and y in terms of R and ϕ where R is the amplitude and ϕ is the phase angle:

$$x = R\cos\phi$$

$$y = R\sin\phi$$

Substituting x and y into Equation (1-1) yields:

$$\rho = R\cos\phi\cos\theta + R\sin\phi\sin\theta$$

Using the trigonometric identity, $\cos(\alpha - \beta) = \cos\alpha\cos\beta + \sin\alpha\sin\beta$, we get:

$$\rho = R\cos(\theta - \phi)$$

Here, R represents the amplitude of the sinusoid, and ϕ represents the phase shift.

Question 2: Parametrization of Lines

We parametrize lines in terms of (ρ, θ) instead of slope-intercept (m, c) because vertical lines cannot be represented in slope-intercept form since the slope would be infinite. Moreover, θ is bounded within $[0, 2\pi]$, and ρ is bounded by the maximum distance from the origin to the image corner, making the parameters more computationally practical. The slope m and intercept c can be expressed in terms of (ρ, θ) as:

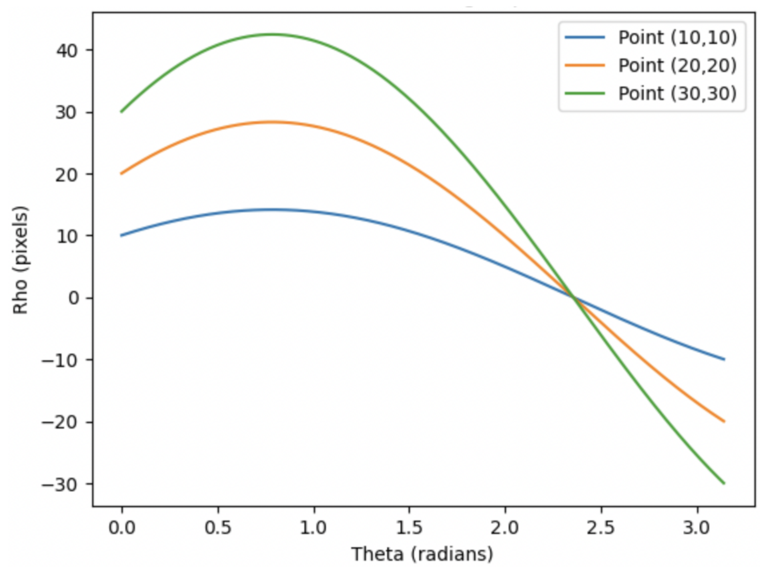
$$m = -\frac{\cos(\theta)}{\sin(\theta)}, \quad c = \frac{\rho}{\sin(\theta)}$$

Question 3: Maximum Absolute Value of ρ and Range for θ

Given $x \in [1, W]$ and $y \in [1, H]$, the maximum absolute value of ρ is the diagonal distance from the origin to the farthest image corner: $\sqrt{W^2 + H^2}$. The range for θ is $[0, 2\pi)$ as it represents the angle of the normal to the line.

Question 4: Plotting Sinusoids and Line Parameters

The sinusoids corresponding to the points $(10, 10)$, $(20, 20)$, and $(30, 30)$ can be plotted in Hough space using Python, and their intersection defines the line through these points: (figure on next page)



To visualize the line detection process in Hough space, we consider three points: (10,10), (20,20), and (30,30). These points are expected to lie on a straight line with the equation $y = mx + c$. In Hough space, this line is represented by sinusoidal curves, where the intersection point of these curves corresponds to the parameters (m, c) of the line in image space.

The intersection point of these sinusoidal waves in Hough space gives the parameters of the line that passes through the three points. Since the points are collinear and follow the equation $y = x$, the slope of the line m is 1, and the y-intercept c is 0.

2 Implementation

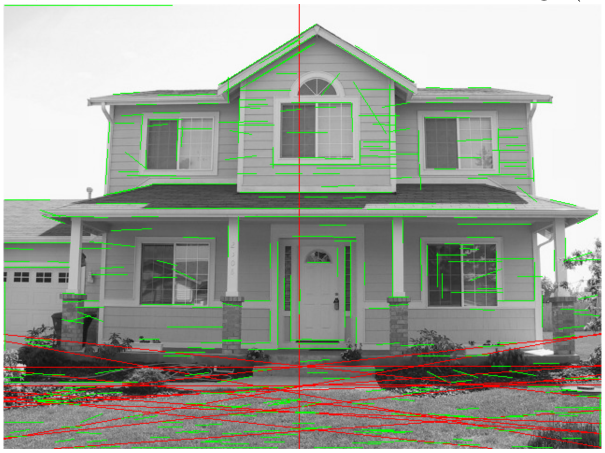
The line detection algorithm involves several stages: image filtering, edge detection, and the application of the Hough Transform for line detection. The performance of the detector was influenced by the choice of parameters at each stage.

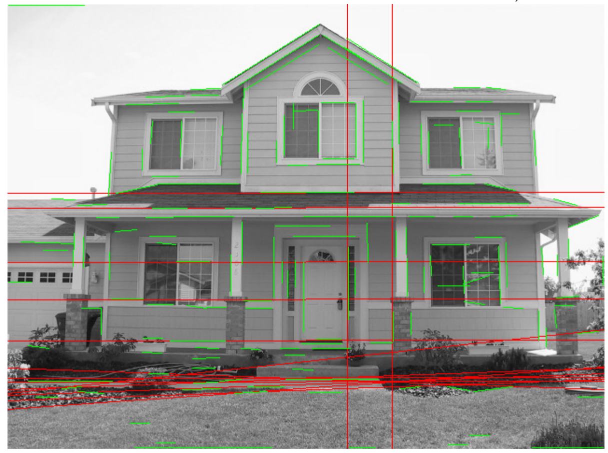
2.1 Image Filtering

The Gaussian filter was applied with a standard deviation of $\sigma = 2$ to smooth the image while preserving edge information.

The choice of σ significantly affects the performance of the subsequent edge detection. For images with a lower amount of noise and finer details (such as the circle or plain cardboard boxes), a smaller σ such as 1 or 1.5 preserved important edge information effectively. However, images with a significant amount of noise or less fine detail (such as the house) benefited from a larger σ, such as 2.5 or 3, which provided better smoothing and facilitated more accurate edge detection.

Below, we can see line detection on the same image (left one with $\sigma = 2$ and the other with $\sigma = 3$.)

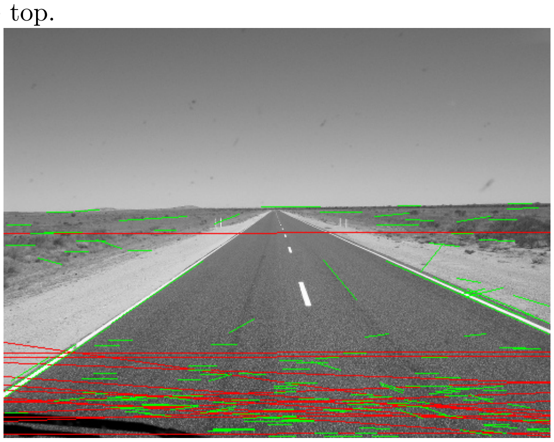


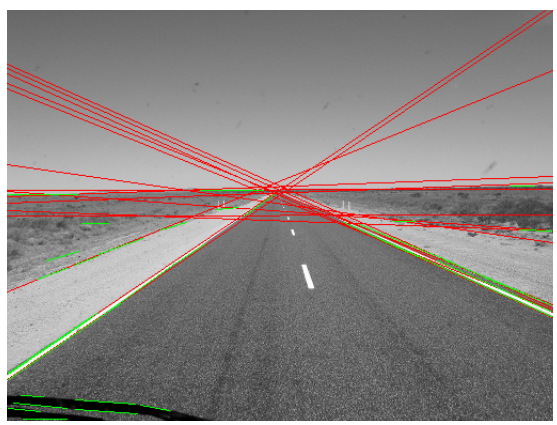


2.2 Edge Detection

The edge detector was given a threshold value of 0.03, which was found to be optimal for the majority of test images. The thresholds for edge detection are critical in determining which gradients in the image are considered as edges. Lower threshold values may include more details but also increase the noise, whereas higher values result in cleaner but potentially incomplete edges. Increasing threshold had a very similar effect (made the line detection much better) on the image as increasing sigma (for noisy images).

Below is an example of a line detection with small threshold of 0.03 (left) contrasted with a big threshold of 0.1. We can see that 0.03 detects too many incomplete edges whereas 0.1 detects a lot of noise as edges on the top.





2.3 Hough Transform

The Hough Transform was applied with a ρ resolution of 2 and a θ resolution of $\frac{\pi}{90}$. The number of lines to be detected was set to 15.

- ρ Resolution: The choice of a 2-pixel resolution for ρ strikes a balance between accuracy and computational efficiency. This resolution is generally sufficient to distinguish between lines that are relatively close to each other while avoiding an excessive number of false positives due to noise or minor variations in the edge map.

- θ Resolution: A θ resolution of $\frac{\pi}{90}$ provides a good level of precision for detecting lines at various angles. This allows the algorithm to detect lines that differ slightly in orientation, which is particularly useful in images with a complex interplay of linear features.

- Number of Lines: Setting the maximum number of detectable lines to 15 helps to ensure that the most prominent lines are captured without overwhelming the image with less significant lines. This is a conservative approach that prioritizes clarity over completeness.

Observations:

For images with finer details or more lines, reducing the ρ resolution and increasing the maximum number of lines might be necessary helped capture the additional detail.

For images with fewer, more distinct lines increasing the ρ resolution improved performance by reducing false positives and focusing on the most significant lines.

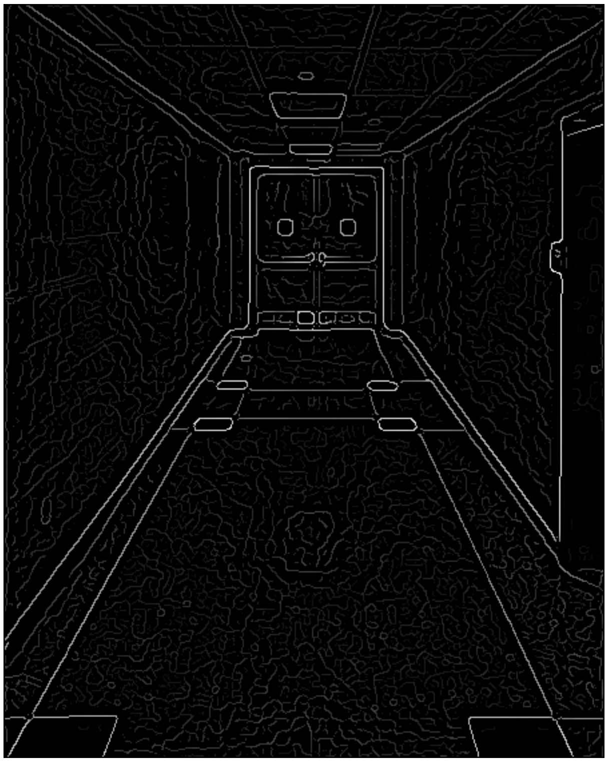
2.4 Results

The line detection process involved several intermediate steps, each contributing to the final outcome. Below are the images that illustrate these steps.

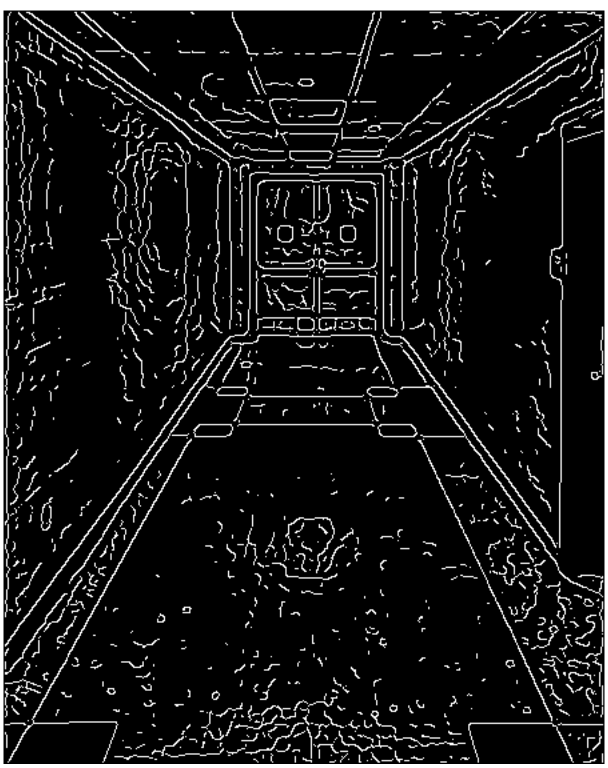
2.4.1 Original Image

2.4.2 myImageFilter(Gaussian Blur)

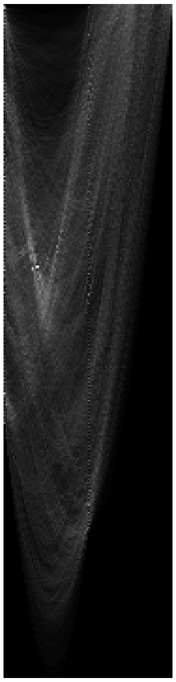
2.4.3 EdgeFilter



2.4.4 EdgeThreshold



2.4.5 Hough Space of Image



2.4.6 Final Line Detection

Submitted by Aryan Dawer on January 25, 2024.