Object Segmentation

1st Vaishnavi V*, 2nd Subhiksha G*,
*Department of CSE, Sri Sivasubramanyia Nadar College of Engineering, Chennai, India
vaishnaviv2410219@ssn.edu.in, subhiksha2410492@ssn.edu.in

***Abstract*—In the domain of image processing, object segmentation plays a major role. COIL-20 dataset is used which is a grayscale dataset consisting of 20 objects captured from 72 different angles with 128x128 pixels, gives a total of 1440 images. The goal of the project is to design a object detection pipeline using only image processing techniques. This includes applying filtering, thresholding, edge detection and morphological operations to segment objects from grayscale images.**

***Index Terms*—Object Segmentation, COIL-20 Dataset, Image Processing, Noise Removal, Edge Detection, Morphology, Contour Extraction, Object Mask**

I. INTRODUCTION

Teaching a computer how to process an image, segment it from its surrounding is a major task in computer vision. This process is called as object segmentation which is the basis for many real world application such as self driving cars, robotics and medical image analysis. Over the past decade, machine learning and deep learning methods emerged into a very popular technique for image segmentation problems. These methods use large neural networks trained by large dataset thus knows how to detect and segment objects automatically. Though these methods gives optimal results, they are called as black box as we can't tell why they resulted this. They have major drawbacks such as they need large training dataset, takes long training time even in powerful hardwares. Thus they become impractical for most of the real world application. Here classical image processing techniques come into picture which does not need large samples of data. This uses mathematical operations directly to pixel values of image. The operations include filtering, thresholding, edge detection, morphological transformations of images. Here, every step is visualised and underrstood unlike machine learning and deep learning. This project builds a complete object segmentation pipeline using only classical image processing techniques applied on COIL-20 (Columbia Object Image Library) dataset. The objects in this dataset include toys, containers, geometric models with a dark background. In summary, this project demonstrates classical image processing without neural networks for object segmentation.

The main contributions of this work are as follows:

- Designing a complete end to end object segmentation pipeline
- Comparison of edge detection methods: Canny & Sobel
- Evaluation of multiple thresholding strategies
- Application of COIL-20 benchmark dataset

II. LITERATURE SURVEY

1) Sobel Gradient Detection: Sobel edge detection is an edge detection technique which uses gradients to identify edges. The Sobel operator, also called Sobel - Feldman operator was introduced as a 3 * 3 isotropic gradient operator for image processing [1]. It contains a pair of convolution kernels, one for vertical and the other, horizontal direction.

The x and y components of the gradient G_x and G_y respectively, are calculated as:

$$G_x = S_x * Z \tag{1}$$

$$G_y = S_y * Z \tag{2}$$

where S_x and S_y are the kernels.

The magnitude of the gradient

$$G = \sqrt{G_x^2 + G_y^2} \tag{3}$$

and the orientation is

$$\theta(x, y) = tan^{-1}(\frac{G_y}{G_x}) \tag{4}$$

The limitation of Sobel is that it produces thick edges and is very sensitive to noise.

2) Canny Edge Detection: Canny edge detection is a multi-stage algorithm developed by John F. Canny in 1986 [2]. It typically consists of 5 stages:

1) Noise Removal: Gaussian filter is applied to smoothen the image in order to remove noise. The formula for Gaussian filter is as follows:
$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \tag{5}$$
2) Gradient Computation: The gradient is calculated using Roberts, Prewitt or Sobel operator.
3) Non-maximum suppression: If there is a neighbouring pixel in the same direction with a higher intensity, then the value of that pixel is suppressed.
4) Double thresholding: Two threshold values (low and high) are fixed. If a pixel value is higher than the high threshold, it is marked as a strong pixel. Similarly, if its value is lesser than the low threshold, then it is a weak pixel and is suppressed. This double thresholding is used to remove pixels caused due to noise.

5) Edge tracking by hysteresis: A weak pixel should be preserved if it comes from a true edge and removed if it is due to noise. For this, a blob analysis is done in which, the 8 neighbours surrounding a weak pixel are considered. If any one of the 8 pixels is a strong pixel, the weak pixel is preserved, otherwise discarded.

3) Hough Transform based Canny (HT-Canny) edge detection: The limitation of the traditional Canny algorithm is that one has to fix thresholds manually but different images may have different lightings. The HT-Canny operator overcomes this limitation. HT-Canny algorithm guided by high threshold image, which obtains edge direction through calculating edge endpoint gradient and connects the edge by using the Hough Transform instead of traditional double threshold method [3].

4) Morphology: Morphology is an image processing technique used for analysing geometry in binary or grayscale images. This is done by probing the image with a set of known shape called structuring element. Some of the morphological operations are dilation, erosion, opening and closing. The erosion of an image all the structures that cannot contain the structuring element and also shrinks the others [4]. The opening operator is used to dilate the eroded image to recover as much as possible the original image. Closing connects the fragmented boundaries, useful for contour extraction [5].

III. Methodology

A. Dataset Organization

1) Dataset Description: The dataset used in this project is COIL-20 processed dataset. COIL-20 is wellknown benchmark dataset in computer vision and image processing, which is widely used for object recognition, segmentation and feature extraction tasks. This contains 20 objects captured from 72 different angles which results in 72 images per object which results in total of 1440 grayscale images.

2) Image Specifications: All the images are processed and captured uniformly spaced angles with resolution of 128x128 and in grayscale format. The background is very clean, processed and stored in PNG format making it easier for researchers to focus mainly on the object.

3) Object Categories: This includes 20 objects which represents a diverse collection of everyday objects with different shapes, sizes, surfaces, textures etc. The object includes toys, household containers and many other items.

B. Pre-processing and Input Construction

1) Noise Reduction: The first step of the image processing pipeline is noise reduction which is used to smoothen the image.

Gaussian filter was used in the traditional Canny and Sobel detections, the formula given by equation (5).

In the HT-Canny algorithm, the median filter was used. The median filter is effective against salt-and-pepper noise also called impulse noise. It replaces each pixel value with the median of its neighbours.

2) Edge Detection: For edge detection, three different methods were used.

- Canny Edge Detection
- Sobel Edge Detection
- Edge Detection using the HT-Canny algorithm

3) Morphology: The morphology operation closing was then applied on the resultant images to connect the fragmented boundaries and to fill small holes and get smooth boundaries. Let $E(x, y)$ be the edge image. Then the closing operation is performed by

$$E \bullet B = (E \oplus B) \ominus B \tag{6}$$

where $\oplus$ is the dilation operator and $\ominus$ is the erosion operator.

4) Contour Extraction and Mask Generation: A set of contours is generated and the largest contour is selected. Let the largest contour be C^*.

Then a filled mask is generated by

$$M(x, y) = \begin{cases} 1, & \text{if } (x, y) \in C^* \\ 0, & \text{otherwise} \end{cases} \tag{7}$$

IV. Results and Discussion

A. Overview

The segmentation pipeline is applied to all 1440 images using two methods: Canny and Sobel. Results were analysed visually.

B. Canny Pipeline

This produced a clean, accurate binary mask in most of the objects. the double thresholding mechanism suppressed the noise while giving strong boundaries.

C. Sobel Pipeline

This produced acceptable results but comparatively lower boundary precision. Here edges were thicker and captured unwanted textures occasionally.

D. Improved Canny (HT-Canny) Pipeline

The HT-Canny algorithm produces most accurate results in segmentation mask among the three. Since it automatically find the threshold values from the image itself, it works well for most of the images. The median filter used before edge detection, removed noise without blurring the images.

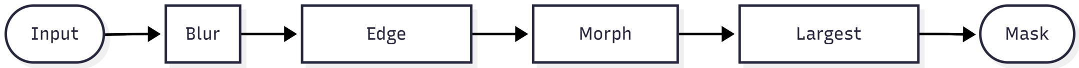

Fig. 1. Flowchart

E. Comparison

Among the three methods, Sobel produced the thickest edges and sometimes captured unwanted textures. Traditional Canny gave better and cleaner edges than Sobel. HT-Canny performed the best overall since its threshold is calculated automatically for each image.

V. Conclusion

A complete object detection pipeline was implemented using only image processing techniques [6]. Edge detection techniques like Canny edge detection and Sobel edge detection were compared. The Sobel based segmentation produced thicker edges compared to Canny. Also, Canny edge detection gave better results for textured surfaces. Additionally, the HT-Canny algorithm, a recently proposed Hough transform based Canny edge detection, was also tested. On comparing all three, the HT-Canny produced the best results. The HT-Canny uses an adaptive threshold which overcomes the limitation of manual threshold of the traditional Canny algorithm. The morphological operation closing was applied to connect fragmented boundaries. Finally, object mask was was generated using the contours.

References

[1] I. Sobel, G. Feldman *et al.*, "A 3x3 isotropic gradient operator for image processing (1968)," *a talk at the Stanford Artificial Intelligence Project*, 1968.

[2] J. Canny, "A computation approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 8, no. 6, pp. 670–700, 1986.

[3] R. Song, Z. Zhang, and H. Liu, "Edge connection based canny edge detection algorithm," *Pattern Recognition and Image Analysis*, vol. 27, no. 4, p. 740–747, Oct. 2017. [Online]. Available: https://doi.org/10.1134/s1054661817040162

[4] P. Soille, *Erosion and Dilation*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 63–103. [Online]. Available: https://doi.org/10.1007/978-3-662-05088-0$_3$

[5] ——, *Opening and Closing*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2004, pp. 105–137. [Online]. Available: https://doi.org/10.1007/978-3-662-05088-0$_4$

[6] Subhiksha and Vaishnavi, "Github - vv-26/segmentation-coil20: Binary object segmentation using image processing." [Online]. Available: https://github.com/vv-26/segmentation-coil20/tree/main

TABLE I
COMPARISON OF CANNY, SOBEL AND HT-CANNY RESULTS

Canny	**Sobel**	**HT-Canny**