Exercise 1 Report

Intelligent Signal Processing - Final Assessment

Overview

This report focuses on discussing the idea behind the methodologies used to implement the solutions for Task 1 and Task 2 of the first exercise. The steps of the implementation are explained using markdown cells in the submitted Jupyter notebooks.

Techniques Used

As directed in the task sheet, the primary technique employed to implement solutions for both tasks is background frame subtraction. Additionally, a selection of pre-processing techniques was used to enhance shape detection within frames.

Background Frame Subtraction

Background frame subtraction is a computer vision is a technique used in computer vision and video processing to identify moving objects within a video sequence by removing the background and isolating the foreground objects. This is achieved by subtracting a background frame from the actual frame, resulting in a foreground mask.

When working with static security camera footage, the simplest approach to obtaining an effective background frame is to select a frame without any moving objects manually. However, this method is not reliable and robust. So in the submitted application, the background frame is calculated by randomly sampling 200 frames from the video and calculating the median frame of all the frames. This approach offers a more dependable and consistent background representation.

Pre-Processing

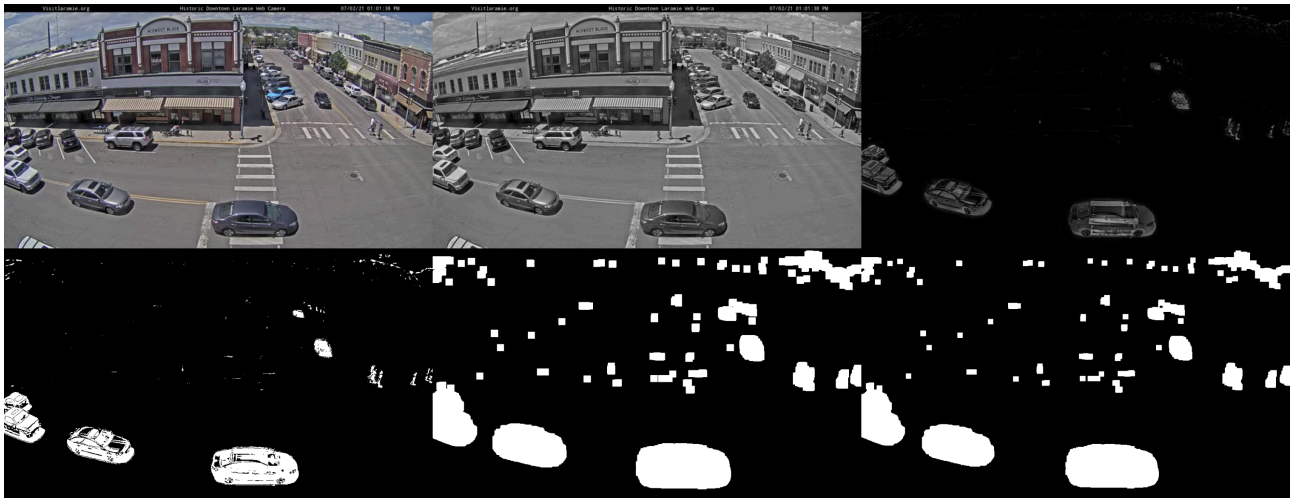

Figure 1: This image displays a frame at various processing stages, following an order from left to right and top to bottom. The sequence begins with the raw frame (top left), followed by the gray frame, the foreground mask, the foreground mask after thresholding, the foreground mask with dilation filtering, and, finally, the foreground mask with erosion filtering (bottom right).

However, though the foreground mask has the background removed, the findContours function from OpenCV will not work well with this since the shapes are not well-defined. To help with shape detection the following pre-processing techniques were applied in a sequence.

Thresholding: Thresholding is an image processing technique used to create binary images from grayscale images. It involves setting pixel values above or below a certain threshold to specific values (usually black or white). This helps separate the objects of interest from the background.

Dilation: Dilation is a morphological operation that expands the boundaries of objects in a binary image. It involves moving a small kernel over the image and replacing each pixel with the maximum pixel value within the neighbourhood defined by the kernel. Dilation is often used to fill gaps, connect disjointed objects, and make objects more prominent.

Erosion: Erosion is the opposite of dilation. It shrinks the boundaries of objects in a binary image. Erosion involves moving a structuring element over the image and replacing each pixel with the minimum pixel value within the neighbourhood defined by the kernel. Erosion is useful for removing small noise, thinning object boundaries, and separating merged objects.

The effects of the pre-processors can be observed in Fig 1.

Results

Task 1

The solution for the first task worked very well and it was able to track all the cars on Main Street. Fig 2 shows a screenshot of the tracking work.

Figure 2: Screenshot of the application built for task 1 tracking cars in the `Traffic_Laramie_1.mp4` video file

Task 2

The solution for the second task was able to successfully count the cars heading towards the City Centre in both the provided videos. The below table presents the results.

File Name	Total Cars	Cars Per Minute
Traffic_Laramie_1.mp4	6	2.022472
Traffic_Laramie_2.mp4	4	2.264151