Social Distancing Detection

Introduction

Code Components

There are two main components to the program: the setup, which only occurs once in the beginning, and the operation, which is a loop that occurs once for each frame of the input video.

When the program begins running, the first frame of the input video is shown to the user. The user then inputs six points with their mouse. The first four points make up a quadrilateral that approximates a rectangle on the ground plane, which will be referred to as the “region of interest” or ROI. The last two points approximate a six-foot distance on the ground.

The purpose of creating a region of interest with the first four mouse points is to solve the issue of camera distortion. Because the camera is filming from an angle, the conversion rate between physical distance on the ground and pixel distance in the image is not constant. In order to solve this problem, the four mouse points are used to warp the region of interest to create a bird’s-eye-view image. EXPAND BY TALKING ABOUT WHICH FUNCTION IS USED AND TRANSFORMATION MATRIX. This new image, shown below, looks distorted and unclear, but its appearance is irrelevant as it will not be shown to the user. What’s important is that in the warped image, the conversion rate between physical distance and pixel distance is now constant.

In order to prove that this works, a small-scale experiment was performed using LEGO figures as human substitutes. Four figures were placed on a piece of graph paper and a rectangle was drawn around them. On the sides of the rectangle, equidistant tick marks were drawn. Images of the figures were taken from angles approximating those of common surveillance footage. This angle caused the tick marks on the side of the rectangle to distort, meaning that the pixel distance between each tick mark decreased as the distance from the camera increased. The pictures were run through the same perspective transformation function described above. The drawn rectangle was used as a guide to input the mouse points. The results of the experiment were successful: the pixel distance between each tick mark on the sides of the rectangle appeared to be equal. This confirmed that the perspective transform function could be used to make the conversion rate between physical distance to pixel distance constant.

The final step of the setup phase is to calculate the minimum safe distance, or the number of pixels that make up six feet once the image is transformed. To do this, the coordinates of the fifth and sixth mouse points are warped using the same transformation matrix used to warp the image. This gives the two points new (x, y) coordinates as they are mapped to the bird’s-eye-view image. The distance formula is then used to calculate the Euclidian distance between these two warped points. The result is a float value that represents the number of pixels that make up a six-foot distance on the ground plane. Because the two points are warped and mapped to the bird’s-eye-view image, this pixel distance is constant throughout the whole image.

The first step in the operation loop is person detection. This is accomplished using a real-time object detection program called You Only Look Once, or YOLOv3. YOLO recognizes a wide variety of objects including people, animals, furniture, vehicles, and many other common items. This specific program included a filter so that only human detections were kept. In addition, one of the arguments passed into the program by the user is a confidence value between zero and one that filters out weak and uncertain detections by YOLO. The default value for this argument is 0.5 but it can be changed using the “--confidence” flag when running the program. Once detection occurs, the results are stored as a list of bounding boxes, or rectangles whose coordinates surround the person being detected. The final part of this step is to apply an algorithm called “non-maximum suppression,” which removes all extraneous boxes to ensure that there is only one box for each person. This algorithm takes in a threshold argument between zero and one that is defaulted to 0.3 but can be changed by the user using the “--threshold” flag when running the program.

PARAGRAPH ABOUT CENTROID TRACKING ALGORITHM

PARAGRAPH ABOUT COORDINATE MAPPING AND DISTANCE CALCULATIONS

PARAGRAPH ABOUT CREATING DISPLAYS

CODE FLOWCHART

Program Usage

Read me info

Next Steps and Conclusion