COMPUTER VISION

To identify the angle of disturbance to the inverted pendulum, the Hough transform feature extraction method in conjunction with canny edge detection method were used. Using the position of the inverted pendulum with no disturbance (i.e.$\theta = 0\ degree$) as our reference line, the angle of disturbance was found to be equal to the angle θ obtained after performing Hough transform and locating the pendulum's position in the image. This relationship was obtained using trigonometric principles as shown below:

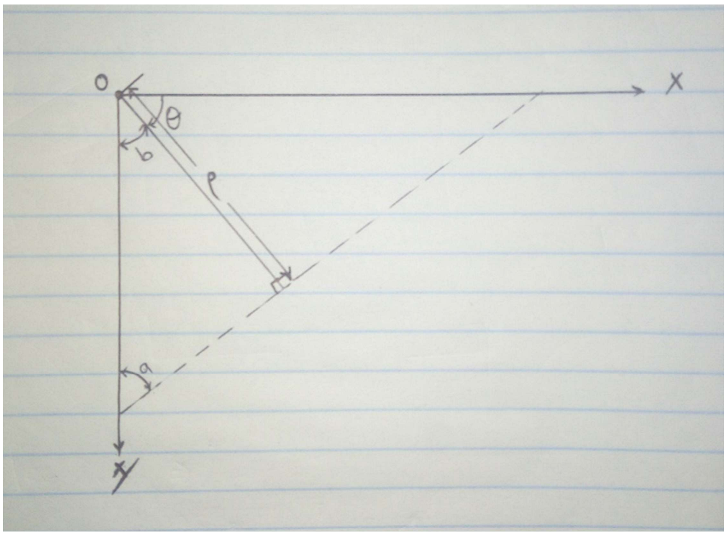


Fig 1: θ *and* ρ *are obtained from the hough transform*

Angle b = $90 - \theta$,therefore from trigonometric properties of a triangle

Angle a = $180 - (90 + 90 - \theta) = \theta$ thus angle of disturbance = θ

The reference line was set such that it aligned with the y axis. This therefore meant that a clockwise disturbance with respect to this reference line corresponds to a positive valued

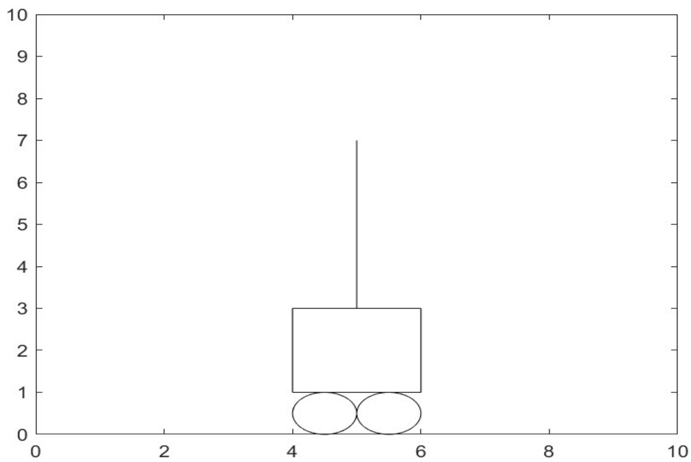


Fig 2: pendulumn at the reference position

rotation and an anticlockwise disturbance corresponds to a negative valued rotation. Thus the first tunable parameter to be set were the pixel coordinates of this reference line.

The reference line was obtained by loading the reference image onto MATLAB, cropping it to a desired size so as to remove the axes data (which were originally included for proportionality purposes of all input frames) and determining the pixel location of the first pixel with a value less than 255 using the ‘find’ functionality in MATLAB. This corresponded to the pixel located at location (4,3). Thus the lower row limit of the inverted pendulumn was found to be row y = 3. Then the find function was used again to find the pixel location of the first pixel with a value less than 255 with the value y = 3 being set as a threshold level so as to detect the first dark pixel for all pixel locations above this threshold value. This was then found to be the pixel value corresponding to location (5,7). Thus the limits of the pendulumn were found to be lower limit

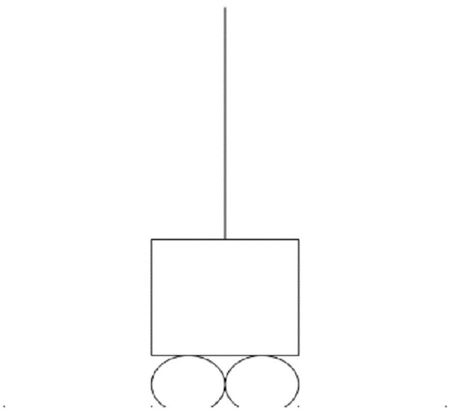


Fig 3a: cropped input without axes data

Fig 3b:detected reference line

(5,3) and upper limit (5,7).

The workspace of the inverted pendulumn was chosen to be within the range -90 to 89 degrees. The reachable workspace of this pendulumn was then divided into two sections the positive disturbance and negative disturbance regions. To determine which region the pendulumn was in, the input frame was first cropped along the line y = 4 and the upper portion was chosen. This was the second tuneable parameter chosen. The value 4 was chosen since some pixel values corrseponding to the line y = 3 had values less than 255 which overspilled into adjacent rows leading to erroneous sectional classification of the lines. An exception to this rule was made in the instance where the angle of disturbance was equal to negative 90 degrees in this case, the threshold was lowered to the line y =3.

On cropping, the line was then classified as being in the positive or negative section. This was achieved by finding the location of the first pixel in the cropped image that had a pixel intensity value of less than 200, i.e the location of the first non-white pixel. If the pixels’ column value was less or greater than the column location corresponding to column x = 5, the image was classified as being in the negative or positive region respectively. In the case where the columns were coincident, the angle of displacement was automatically set to zero.

The pixel coordinates corresponding to the reference line occupied two columns. The third tuneable parameter chosen was the lower of the two columns indices. In this particular case, the reference line was found to lie between columns 298 and 299. The 298^{th} row was chosen as the reference line since pixels found at column locations less than 298 were in the negative region creating a clear distinction between the two regions.

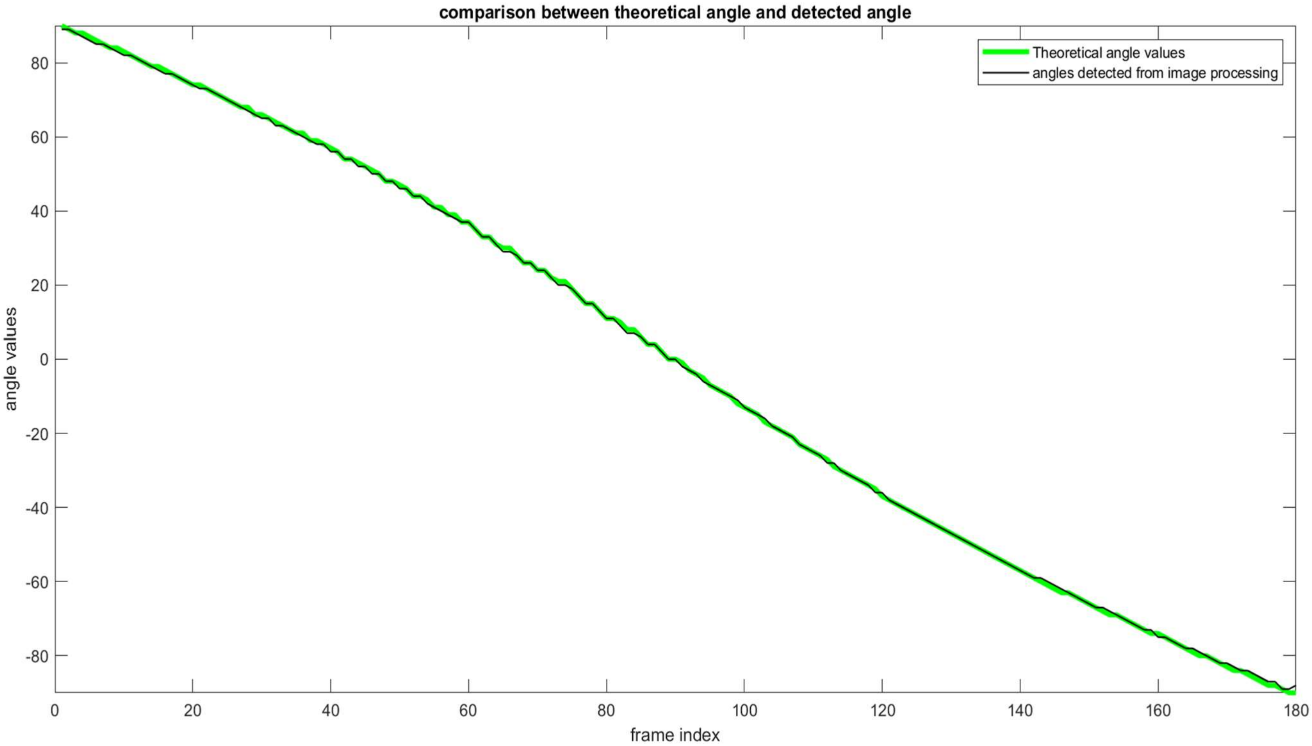
Finally, angles which lay in the negative quadrant were then treated as a reflection of an equivalent angle in the positive quadrant. This meant that in the case where the pendulumn was detected to be in the negative quadrant, the image is flipped, analyzed and the resultant angle is then negated to indicate that it is a counter clockwise disturbance of the inverted pendulumn.

To achieve the most accurate results, the pixel coordinates of the pendulum's point of rotation need to be accurately determined. These values are used in calculating the actual angle of disturbance of the pendulum (using trigonometric principles) which is then compared against the angle obtained from image processing so as to check the accuracy of the image processing algorithm.

Also the use of an edge dection algorithm, for example canny edge detection in conjuncction with the hough transform reduces the computational cost required in detecting the straight line. This is the case since applying the hough transform without edge detection results in a a large number of votes in the parameter space for every distinct point in the input image. Thus for a 256 by 256 image the (m, c) parametrization requires 256 votes to be accumulated, while the (θ, ρ) parametrization requires a similar number. (Davies, 2012)

The number of votes is even higher (360) if the θ quantization is to be fine enough to resolve 1 degree changes in line orientation (Davies, 2012) . With the use of canny edge detection, Hough transform analysis is only performed on the detected edges thus significantly reducing the number of distinct points in the input image which consequently reduces the computational cost involved.

The main advantage of the computer vision in this task is that it's accurate. Using the case of the video input to the algorithm, a comparison between the calculated theoretical angle and angle detected from image processing is as show below:



The main limitation of computer vision technique is that it is custom built to address a particular problem therefore limiting its applications. In this case, the system is designed to detect disturbances in an ideal world where there is a distinct colour difference between the foreground and background and the reference line is set with respect to the size of the image and location of the inverted pendulum. They are also resource intensive since by nature, still images and video frames are large therefore incurring high computational costs in this case the input video file is 6 seconds long and is made up of 180 frames each with the dimensions 420 by 560, making it a total of 42,336,000 pixels.

Reference

1. Davies, E. R., 2012. *Computer and Machine Vision - Theory, Algorithms, Practicalities.* 4th ed. Waltham: Elsevier Inc Academic Press.