CS4053: Assignment 3

Submitted by-Yash Pandey

17317629

D-stream

Objective:

To develop a computer vision solution to locate if and where objects are abandoned and/or removed within a video taken by a static camera. Two distinct steps are to be considered while developing the solution:

1) Detection of changes to the static background due to object abandonment or removal.
2) Distinguishing between object abandonment and object removal events.

A combination of background subtraction and object tracking is used to look for certain defined patters of activity that occurs when an object is abandoned in an abandoned object detection system.

Most of the techniques of detection of an abandoned object employ an approach with several independent steps where the output of previous steps serves as an input to next step. These steps are as follows:

1) Background Modelling and subtraction: a dynamic model of the background is created and subtracted from each incoming frame to detect the current foreground. The output is a mask depicting pixels in the current frame that do not match the background.
2) Foreground analysis: this step is used to remove any unnecessary false foreground regions that may be caused by factors such as lightening variations, etc and to correct shadows.
3) Blob extraction: foreground blobs are detected and blobs that are too small and caused by noises are discarded.
4) Blob tracking: in this step a correspondence is found between the current foreground blobs and the existing tracked blobs from the previous frame if any.
5) Abandonment analysis: in this step a static blob detected bin the previous step is classified as either abandoned or removed object.

Solution:

1) Pre-processing-

a) Contrast enhancement:

In this step the quality of low light videos are improved by normalising the difference between minimum and maximum intensities in the image which helps to increase the visibility in the darker regions. The method that can be used is Histogram Equalisation, in which the input RGB is split into three greyscales one corresponding to each channel after which the histogram of each greyscale is computed and normalised so that the sum of histogram bins become 255. Image corresponding to the transformed histograms is computed using their integral after which the 3 transformed greyscales are joined to get the output RGB image.

b) Noise reduction:

The white noise present in the input frame is reduced by smoothing the frame. This can be done by subjecting the image to one of the several smoothing filters such as gaussian blurring filter.

2) Background modelling and subtraction:

Adaptive median can be used, it is assumed that the background is more likely to appear at any given pixel over a period than the foreground objects. This method is relatively easy to perform from computational point of view, but it has high memory requirements as the previous frames must be stored in a buffer at any given time and the number must be quite large to get a good estimate of actual background. The pixel stays in the background for majority of the values and hence the median of the previous n frames can be used as the background model. The running estimate of the median is incremented by one if the current intensity is larger than the existing estimate and it is decreased by one if it is smaller. The estimate is left unchanged if it is equal to the current pixel value.

$Bt = ((B_{t-1}-1)\ if\ B_{t-1} > It;\ (Bt-1)\ if\ B_{t-1} = It; (\ B_{t-1}+1)\ if\ B_{t-1} < It)$

B_t is intensity of background model

L_t is the intensity of the current model

3) Foreground Analysis:

Shadows can be detected using the single step Complex NCC method in which NCC is calculated for each pixel using the following expression:

$$NCC = \frac{\Sigma_{u \in W} I_f . I_b \frac{1}{MN} \Sigma_{u \in W} I_f \, \Sigma_{u \in W} I_b}{\sqrt{(\Sigma_{u \in W} I_f^2 \frac{1}{MN} (\Sigma_{u \in W} I_f)^2)(\Sigma_{u \in W} I_b^2 \frac{1}{MN} (\Sigma_{u \in W} I_b)^2)}}$$

W is MxN neighbourhood centred at that pixel

I_f is the current frame

I_b is the background intensity values at a pixel

A pixel is classified as shadow if NCC>= T_{shadow}

4) Blob Extraction:

An algorithm can be used to detect external as well as internal contours and label the internal points of each connected components.

5) Blob Tracking:

Blob correspondence is to be established. Each blob in the incoming frame is compared with the existing blobs in the tracking system to find a match based on position and size Euclidean distance between the centroids of the two blobs can be measured and if the distance between positions is less than a specified fraction of the length of its bounding box diagonal then existing blob will match the new blob. Once the positions of the two blobs match then size is compared by the bounding box areas. Miss count and hit count are also maintained.

6) Removed object detection:

Edge detection can be used to analyse removed objects. If we place an object in the front of the background it will introduce more edges to the scene around the object boundaries. The points located along the object's external contours are checked for high intensity values in the gradient images obtain the edge energy associated with that object. The object is classified as removed if this energy is greater for the background image and if not then its considered abandoned.