Pedestrian Detection using Background Subtraction and Histogram of Oriented Gradients

Group-CS23

Mentored By : Rajitha B.

17th April 2017

Introduction

Problem Statement : Efficient human detection using Background Subtraction and Histogram of Oriented Gradients.

Problems generally encountered :

- Dynamic nature of the Background.
- Non periodic Human Motion.
- Occlusions and large number of possible appearances.

Overview of Procedure

Pedestrian Detection process can be divided into three stages:

- Region of Interest Selection i.e Motion Detection.
- Selection of effective features.
- Classification whether a Pedestrian or not.

Overview Of Proposed Approach

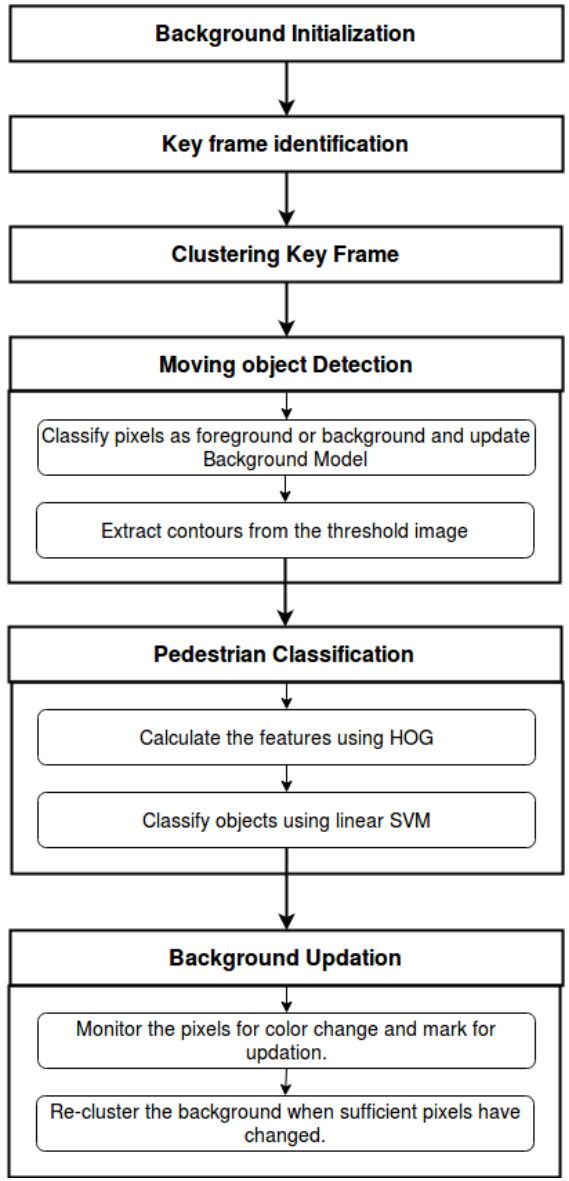

Figure 1: Flow of Algorithm

Background Modeling

Efficient Background Modeling is required for motion detection.
Using this model we separate foreground objects.
Problems generally faced :

- Dynamic texture scenes like waving leaves,rippling water.
- Lighting - Shadows and Gradual/Sudden illumination variations.
- Color based differences and noise.
- Foreground object becomes motionless and a part of the background eventually.

Key Frame Identification

The procedure followed is :

- Convert the current frame to gray scale.
- Take the absolute intensity difference between current gray frame and previous frame.
- If the difference is greater than given threshold it is a Key Frame.
- This key frame is further passed for Fuzzy_Predict Cluster analysis.

Clustering

A loose definition of clustering could be **"the process of organizing objects into groups whose members are similar in some way"**.

- Hard clustering method means partitioning the data into a specified number of mutually exclusive subsets and requires that an object either does or does not belong to a cluster.
- Fuzzy C-means is an overlapping clustering algorithm. It is more natural than hard clustering.
- Objects on the boundaries between several classes are not forced to fully belong to one of the classes.

Fuzzy CMeans Clustering

In fuzzy clustering, each point has a degree of belonging to clusters, as in fuzzy logic, rather than belonging completely to just one cluster.

- x - vector of image pixel intensities
- $u_k(x)$: the degree of being in the kth cluster
- c - vector of cluster centroids
- m - fuzziness degree
- N - number of pixels in image
- K - number of cluster centroids

$$c_j = \frac{\sum_{i=1}^{N} u_{ij}^m . x_i}{\sum_{i=1}^{N} u_{ij}^m} \qquad u_{ij} = \frac{1}{\sum_{k=1}^{C} \frac{||x_i - c_j||}{||xi - c_k||}^{\frac{2}{m-1}}}$$

Flowchart of Fuzzy Clustering

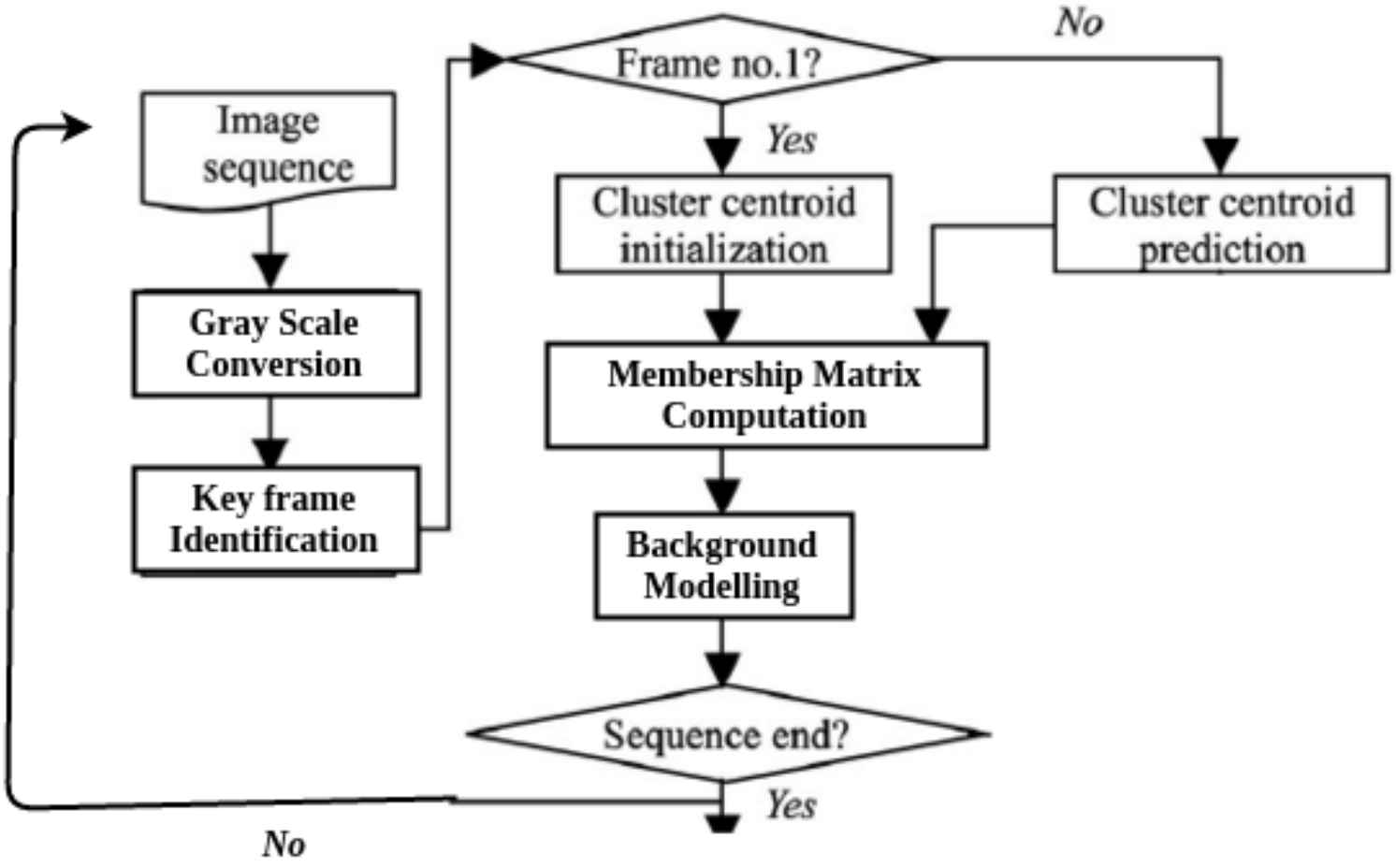

Figure 2: Flow of Fuzzy cmeans Algorithm

Clustering Example

Figure 3:
Frame

Figure 4:
Clusters K=2

Figure 5:
Clusters K=3

Figure 6:
Clusters K=4

Advantages and Disadvantages

Advantages:

- Gives best result for overlapped data set and comparatively better then k-means algorithm.
- Eliminates noisy spots and reduces false blobs.
- Less sensitive to noise and more homogeneous regions are obtained.

Disadvantages:

- Large Computational time.
- Prior specification of the number of clusters.
- Sensitive to initialization condition of cluster numbers and cluster centers.

Foreground Detection and Update Background Model

- Every pixel is classified as either foreground (255) or background (0) depending on the similarity measure ρ.
- For every Background pixel, the Reference membership matrix u_i is updated using learning rate α.

Similarity ρ : $\rho(u_i^b, u_i) = \sum_{j=0}^{c-1} min(u_j^b, u_j)$

Update Background Model : $u_i^b = (1 - \alpha) * u_i^b + \alpha * u_i$

Foreground Detection and Contour Extraction

Figure 7: Foreground Object Detection

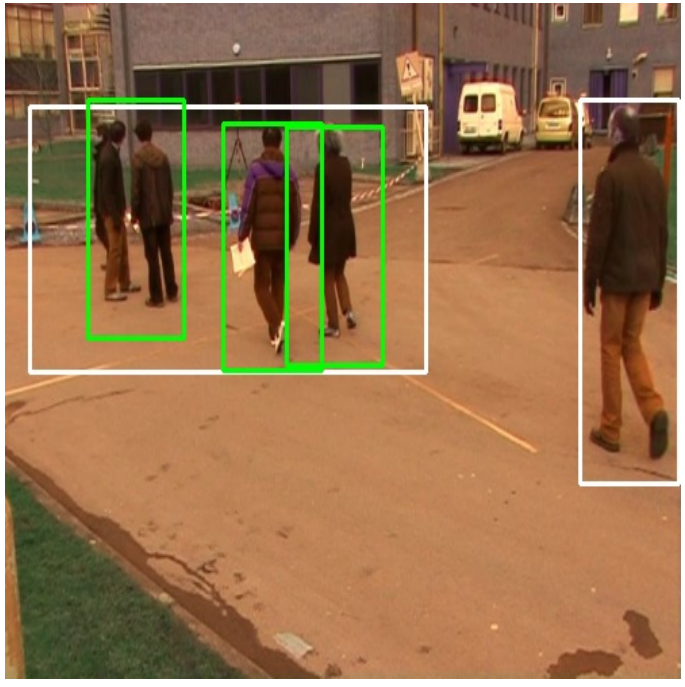

Figure 8: Extraction of moving objects

Previous Approaches For Human Detection

Human Detection Techniques :

- Using Edge Templates by Gravila Et Al in 1998.
- Haar wavelets powered by Adaboost given by Viola and Jones in 2001.
- Key Point Detectors on backgrounds in 2004-2005.

Histogram of Oriented Gradients

After motion detection, sliding windows can be used to shift on detected regions.

- Each window means an individual image region, and corresponding HOG feature vector will be calculated, which reflect the edge and gradient information of image region.
- The basic idea is the local object appearance and shape can often be characterized rather well by the distribution of local intensity gradients or edge directions.
- The algorithm is based on evaluation of well-normalized local histograms of image gradient orientations in a dense grid.

HOG Gradient Computation

Figure 9: HOG Gradients

HOG Working

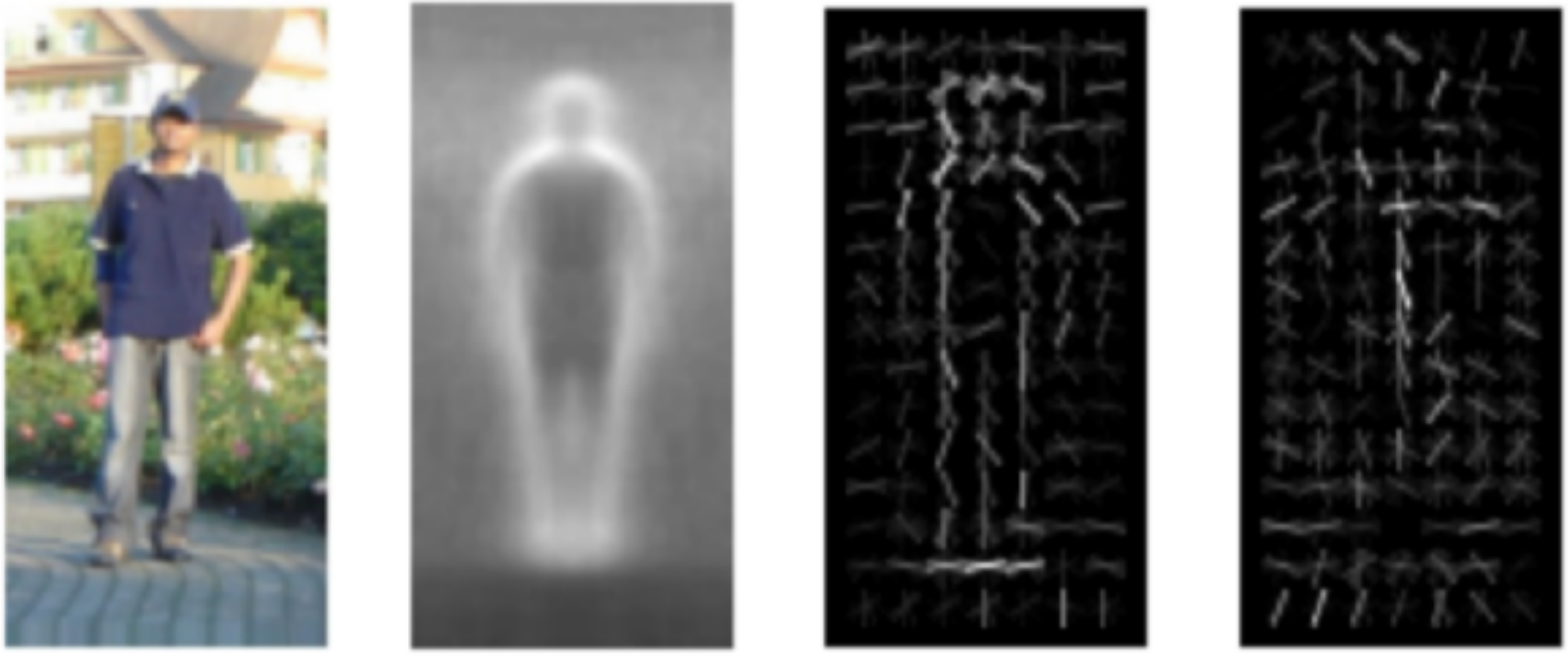

Figure 10: Working

HOG Implementation

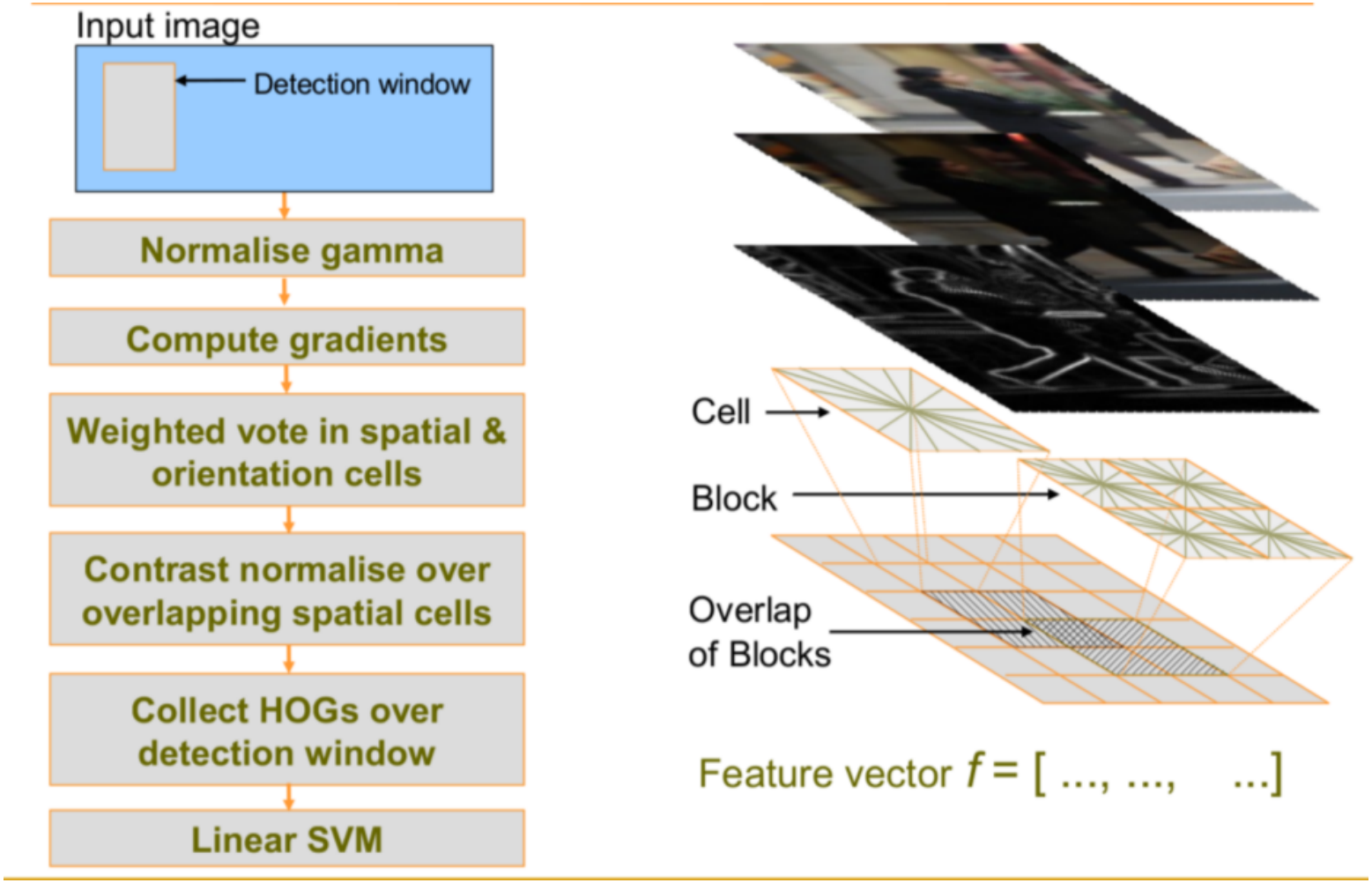

Figure 11: Process Flow

HOG Algorithm

Basic steps to implement HOG :

1. The color image is converted to grayscale.
2. The gradient is calculated at each pixel.
3. To create a histogram of gradient orientations for each cell.
   Feature quantity becomes robust to changes of form
4. Normalization and Descriptor Blocks :
   Feature quantity becomes robust to changes in illumination.

HOG Computation

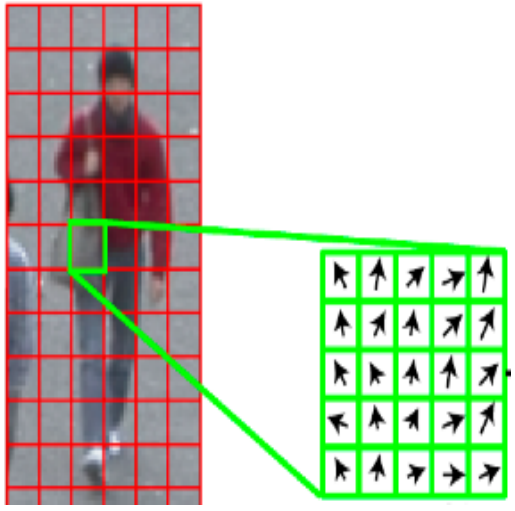

Figure 12: Gradients in a cell

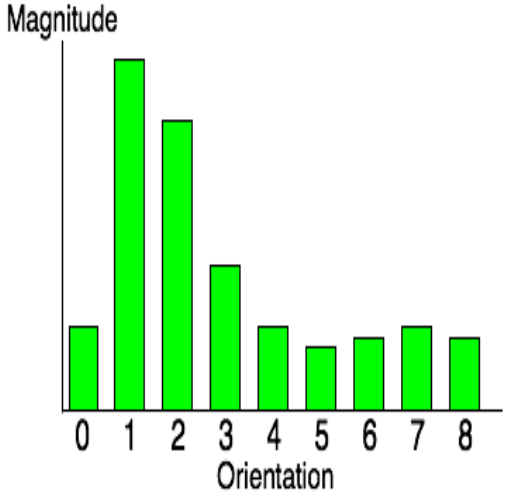

Figure 13: Orientation Binning

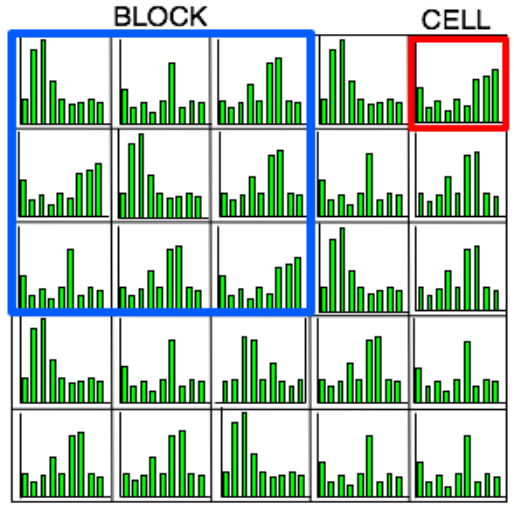

Figure 14: Extracted HOG

Classification: Support Vector Machines(SVM)

- Linear SVM is the classifier which classifies the detected object as human or non-human.
- SVM classifier should be trained with the positive samples and the negative samples of the sample data before the classifications.
- The sample data is the HOG feature vectors of general human and background images.
- When the support vectors are trained sufficiently, the linear SVM classification can be used to recognize humans from static images or real-time video.

Linear SVM

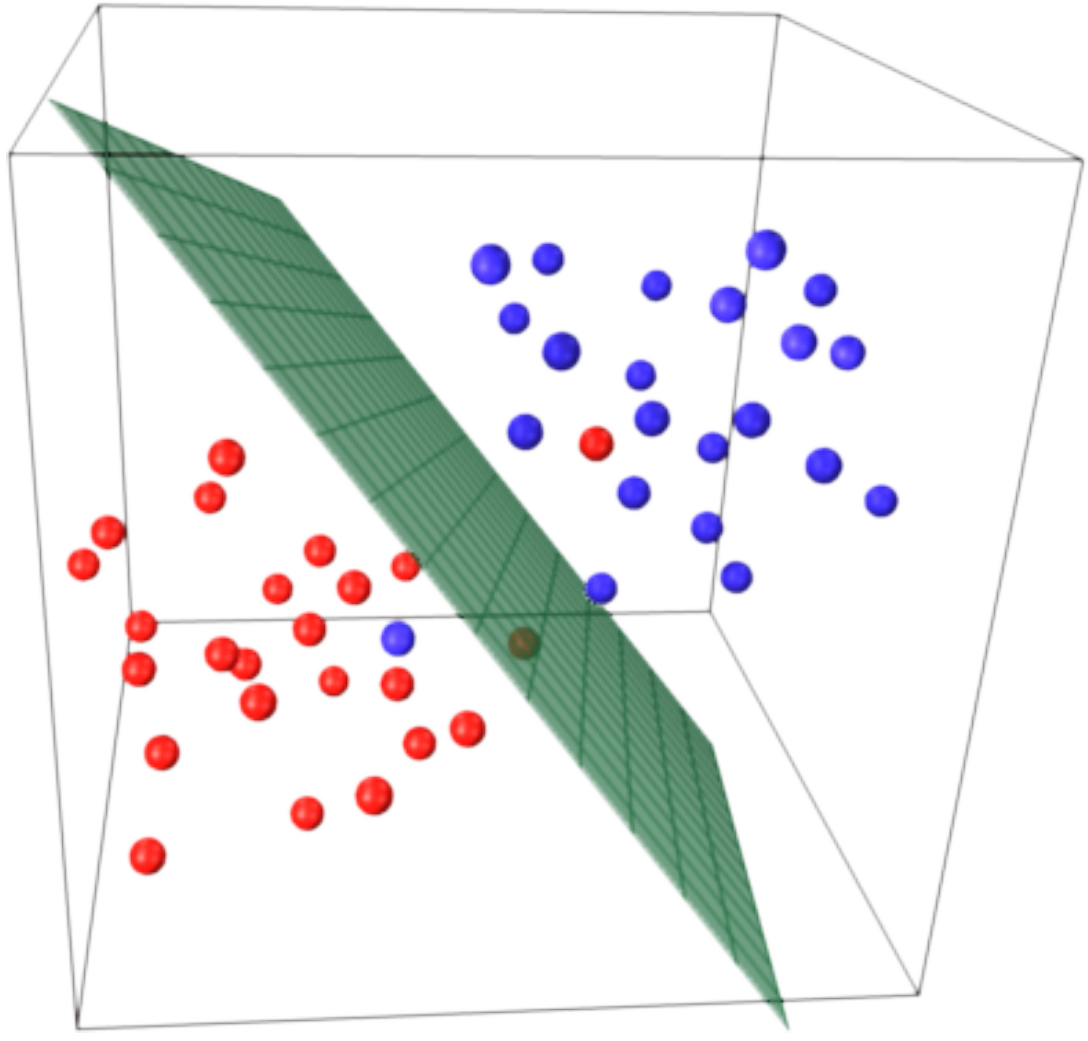

Figure 15: Classification

Update Background

Issue : For the objects which enter the frame and become static need to be considered as a part of the background. Thus , the reference background frame needs to be updated and re-clustered.

Steps :

1. **Pixel Colour Monitoring** : The pixels in contours which have been rejected by HOG are monitored for their colour change. If the new colour appears continuously over *bg_colour_freq* frames then the pixel is marked for updation.

2. **Background Re-clustering** : When the number of pixels marked for updation exceeds *bg_cluster_thresh* then the updated background image is re-clustered using Fuzzy cmeans algorithm.

Background Updation Results

Figure 16: Initial Background

Figure 17: Updated Background

Result Analysis

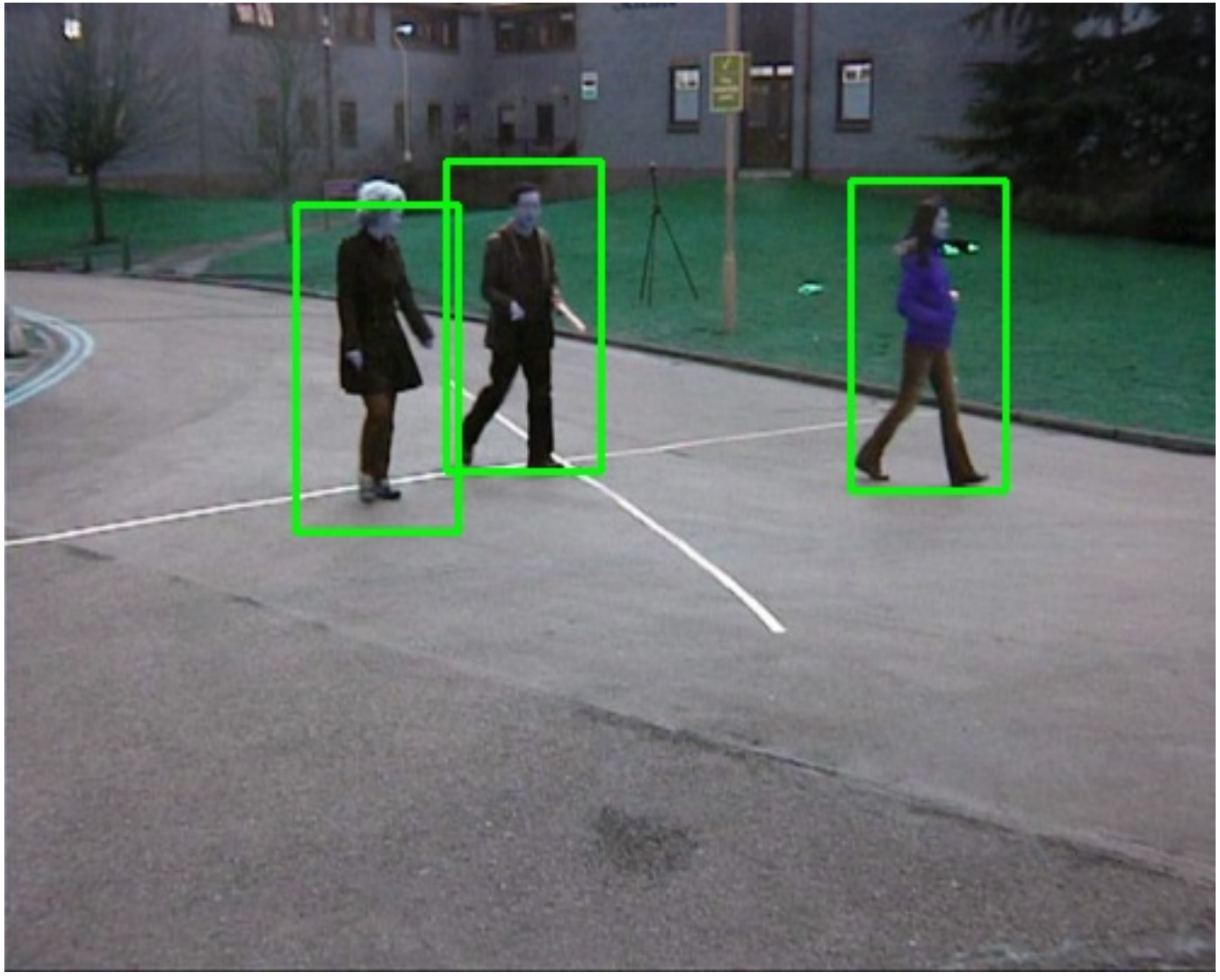

Figure 18: Pedestrian Detected

Results for PETS Datasets

Fuzzy and HOG	Datasets			
	A	B	C	D
True Positive	229	226	972	338
True Negative	5	0	0	0
False Positive	7	0	0	0
False Negative	30	18	110	48
Total No of Frames	100	50	250	100
Total No of People	259	244	1082	386
Accuracy(%)	86.34	92.60	89.83	87.56

Figure 19: Fuzzy and Hog results

Average Accuracy : 89.08%

Fuzzy	Datasets			
	A	B	C	D
True Positive	219	230	531	360
True Negative	0	0	0	0
False Positive	0	0	0	0
False Negative	10	0	27	26
Total No of Frames	60	50	130	80
Total No of People	229	230	558	386
Accuracy(%)	95.63	100	95.16	93.26

Figure 20: Fuzzy results

Average Accuracy : 96.01%

HOG	Datasets			
	A	B	C	D
True Positive	227	178	469	257
True Negative	0	0	0	0
False Positive	23	39	117	14
False Negative	41	30	265	100
Total No of Frames	60	50	160	80
Total No of People	268	208	734	357
Accuracy(%)	78.00	72.06	55.11	69.27

Figure 21: Hog results

Average Accuracy : 68.61%

Conclusion

- Used Fuzzy Cmeans Clustering to eliminate the problems related to dynamic nature of background.
- We extracted the area where motion is detected through Foreground segmentation and we passed only this area for HOG computation.
- Hence the false positive detections were minimized, improving the accuracy.
- Updated the reference background image.

Thank You!