ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Object Detection and Classification Using Scale Invariant Feature Transform For SMS Reporting System

Babu.p
M.Tech (DECS),
Electronics and Communication Engineering
Annamacharya Institute of Technology and Sciences,
Rajampet, Andhra Pradesh

Lakshmi Devi.P, B.Tech, M.Tech
Associate professor,
Electronics and Communication Engineering
Annamacharya Institute of Technology and Sciences,
Rajampet, Andhra Pradesh

Abstract:

Object detection and classification are two key tasks for image understanding, and have attracted much attention. We propose efficient algorithm for moving object detection, classification and evaluate its parameter by alternating the algorithm in effective way. In proposed system image subtraction, threshold and foreground detection algorithms will be used for object detection and scale invariant feature transform (SIFT) are used for classification. Then frame by frame the objects are tracked and parameters like velocity of motioned object will be calculated. It is mainly developed for protection and monitoring, sports training. With the help of MATLAB here we are reporting the object details by SMS using GSM.

I.Introduction:

Moving object detection in real time application is a important challenging task in video surveillance systems. There are three steps for automated video analysis. They are object detection, object tracking and classification. As the first step, object detection aims to locate and segment interesting objects in a video. Then, such objects can be tracked from frame to frame, and the tracked object can be analyzed to recognize object categorizes.

Thus object detection in video acts as an first step for next processing such as tracking, classification of the detected moving object. Thus, object detection plays a critical role in practical applications and challenging task in real time. The use of object detection, tracking and classification algorithms are not restricted in video surveillance.

The major application of object detection and classifications are broader monitoring and protection, sports training, traffic control etc. Thus object detection play major role in various fields. In most of the work, Optical flow method is used for detecting moving objects in video and image sequences.

For example, adaptive optical flow method for tracking a person is dependent on being able to detect a person in particular across a series of frames in videos. Optical flow can be used to segment and extract a moving object from a scene, provided that estimated velocity of the moving object is known. But successful object detection in video also relies on being able to segment the background from foreground.

However, the high computational time is needed to extract the object in optical flow and the lack of classification the foreground from the background. So this method is unsuitable for real time processing and misclassifies unmoved objects or large texture less regions as background. Then Background subtraction is commonly used technique for segmentation of motion object in static scenes.

It will attempts to detect moving object regions by subtracting the current image in each pixel-by-pixel from a reference background image that is created by averaging images over time period. Unfortunately, the derivation of background subtraction model is Therefore, most of existing approaches for moving object detection and classification are computationally difficult and subject to more delays, adversely affect the performance of real-time surveillance.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

II. Overview of the System:

In proposed system the aim is to build robust and novel moving object detection, classification algorithm that can detect object in a variety of challenging real world scenarios. The Overall system overview will be represented in figure 1.

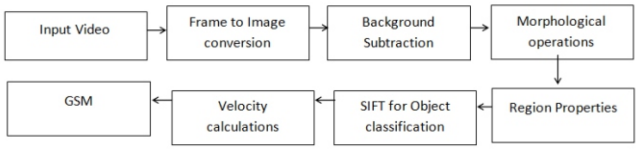


Fig.1 Overview of the system.

The steps involved in overall system is

1. The first step is the acquired video is uploaded in application. The input video will be taken for both static and dynamic background. It will acquire from static camera. For processing the video files, convert video into frames and from frames to images.

2. Next step is by using image subtraction techniques, the difference between two images are estimated for evaluating the movement changes of an object.

3. Then segmentation is performed and foreground object is detected. While foreground detection the unwanted backgrounds are filtered using morphological filtering.

4. After the object detection, the next step is to identify object by comparing with datasets. SIFT is used to extract the features of object and matching will helps to classify the objects in video frame.

5. Then next one is parameters like speed, velocities are evaluated by tracking the object are evaluated by using BW labels.

6. Next the information of object speed and object name will send to authorized user for reporting using serial communication through MATLAB and GSM.

GSM MODULE:

FIG: GSM MODULE

The physical connection between the CPU and GSM module through RS232 cable.Sim card insert the sim slot and power supply given the through charger . The GSM module can be used for counted number of objects presented in videos and their speed, velocity related information passes to administrator through utilizes the SMS services.

FEATURES:

• SIM 300 GSM module used.

• It provides the RS232 serial interface.

• Provides serial TTL interface for easily and directly interface to MATLAB software through external interface.

• GSM can be used for voice communications like SMS, GPRS/GPS, data/fax.

• Can be controlled by standard AT commands.

• 7-15 AC or DC voltage is required.

• Tri-Band GSM/ 900/1800/1900MHz or Quad-Band GSM/ 850/900/1800/1900MHz

• GPRS multi-slot class 10/8

• GPRS mobile station class B

• Compliant to GSM phase 2/2+ Class 4 (2W @ 850/900 MHz) | Class 1 (1W @ 1800/1900MHz)

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

- Control via AT commands (GSM 07.07, 07.05)

BASE BOARD SPECIFICATIONS:

- On-Board Voltage Regulator | Signal Indicators

DB9 Con for Host or MATLAB interface, control by AT commands

SPECIFICATIONS FOR DATA:

- PBCCH support
- GPRS class 10 :max. 85.6 Kbps (downlink)
- PPP-stack | Coding schemes CS 1, 2, 3, 4
- CSD up to 14.4 kbps | USSD
- Non transparent mode

SPECIFICATIONS FOR SMS VIA GSM:

- Point-to-point MO and MT
- SMS cell broadcast | Text and PDU mode

SPECIFICATIONS:

- Iinterface:RS232
- Voltage:50v

II.Moving Object Detection:

Moving object detection is initial step in video surveillances. There are various algorithms used for detecting moving objects. The input video is taken using a static camera. It will be the extension of (E.g.: .Avi) format. It will acquire in both static and dynamic background. The challenges like view point variation, shape, occlusion etc are probably facing.

A. Frame Conversion:

For processing an Input Video files, it has to convert it into frames by finding the information about .avi file. After that it has to convert into images. So videos are split into frames. The first frame, which is called reference frame, which represents the reference pixel values for comparing purpose and the second frame which is called the input frame, which contains the moving object.

The two frames are compared and the differences in pixel values are determined. For each naming is assigned.

B. Difference Image:

The difference between two frames is done with the Difference image techniques. It is nothing but subtraction of images acquired at different instants. That is by subtracting current frame and previous frame for detecting object from background in motion states. A difference image d(i,j) is binary image where non-zero values represent image area with motion, that is, areas where there was a substantial difference between consecutive images f1 and f2:

$$d(i,j) = \begin{cases} 0 & \text{if } |f_1(i,j) - f_2(i,j) <= T| \\ 1 & \text{Otherwise} \end{cases} \quad (1)$$

Where T is small positive number and i,j represent the pixel value of images in row and column wise.

C. Threshold :

The object detection will performed based on thresholding concept. Threshold is one of the segmentation process. It can be determined to segment objects and background. It is the process of separating an image into different portions of images by selecting a certain grayness level as a threshold and comparing each pixel value with that threshold, and then assign the pixel to the different portions, depending on whether the pixel's grayness level is below the threshold or above the threshold value. So threshold is applied and the pixel is identified as background or foreground. A complete segmentation of an image R is a finite set of region R1, R2,………… Rs.

$$R = \bigcup_{i=1}^{S} R_i \quad (2)$$

Thresholding is the transformation of an input image f to an output (segmented) binary image g as follows

$$g(i,j) = \begin{cases} 1 & \text{for } f(i,j) >= T \\ 0 & \text{for } f(i,j) < T \end{cases} \quad (3)$$

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Where T is the threshold, g(i,j) =1 for image elements of object and g(i,j) =0 for image element of background. If the difference average pixel value is smaller than a certain threshold value, then the output image background will be white (that is pixel value is 255). Otherwise the background will be black (pixel value is 0). Now the pixels are classified and the motion is tracked in the video.

Algorithm:

1) Search all the pixels f(i,j) of the image f.

2) Set the threshold value based on image subtraction pixel value.

3) An image element g(i,j) of the segmented image is an object pixel if f(i,j)>= T and is a background pixel otherwise.

4) Extract the foreground object without background.

Finally the segmented objects are extracted without background. Suppose in foreground detection, if background is consider as foreground means then morphological filtering is used where erosion and dilation are performed to minimize the effects of noise and improve the detected regions.

Erosion:

Erosion combines two sets using vector subtraction operation of set elements and is the dual operator of dilation. Neither erosion nor dilation is an invertible transformation.

$$X\Theta B = \{P \in Z^2 : p + b \in X \ \ for\ every\ b \in B\} \quad (4)$$

This formula says that every point p from the image is tested.The result of the erosion is given by those points p for which all possible p+b are in X.

Dilation :

Dilation combines two sets using vector addition (or Minkowski set addition, e.g., (a,b)+(c,d) = (a+c, b+d)). The dilation XB is the point set of all possible vector addition of pairs of elements, one from each of the sets X and B.

$$X \oplus B = \{P \in Z^2 : x + b, x \in X \ \ for\ every\ b \in B\} \quad (5)$$

After the object detection in video the accuracy measures done for output performance. For that the ground truth images are maintained in separate folder for measuring the accuracy. The sample ground truth images are,

Fig 2. Example for ground truth images.

Then the measures like Root Mean Square Error(RMSE) and Peak Signal to Noise Ration(PSNR) are measured. The formula to measure RMSE as,

$$RMSE = \frac{1}{MN}\sqrt{\sum\sum(a-b)^2} \quad (6)$$

The formula to measure PSNR as,

$$PSNR = 10.\log(\frac{255^2}{RMSE}) \quad (7)$$

Where a represents a segmented image and b represents ground truth images and r and c are size of the segmented images. By comparing the segmented images with ground truth image the sensitivity measure is evaluated in percent. If the sensitivity measure is low then the performance of detected object is minimum or else the performance is good.

IV. Object Classifications:

In video analysis the second step is object classification. In proposed work the detected object is classified as whether it may be human being or not. It will base on pattern matching techniques. Patterns are nothing but arrangement of descriptors with certain properties.

Scale Invariant Feature Transform (SIFT) was proposed byDavid Lowe [10] that is able to detect and describe local image features efficiently. The main SIFT algorithm consists of four major stages:

- Key point localization.
- Orientation assignment.
- Key point descriptor.

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

The following sub-section will describe each stage.

1.Key point localization:

The first stage of SIFT is the detection of local interest points called keypoint. In this stage, the algorithm must search the potential keypoints over all scales and image locations. It can be efficiently implemented by using a difference-of-Gaussian function that are invariant to scale and orientation. The scale space of an image is defined as a function that is produced from the convolution of a variable-scale Gaussian G(x, y,σ), with an input image I(x, y)as shown in Eq.8 and Eq.9:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \qquad (8)$$

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \qquad (9)$$

To effectively detect stable keypoint location in scale space, Lowe [17] used scale space peaks in Difference of Gaussian (DoG) function convolved with the image D(x, y,σ) which can be computed from the difference of two nearby scaled images separated by a multiplicative factor k as in Eq. 10:

$$D(x, y, \sigma) = \begin{matrix}(G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) \\ L(x, y, k\sigma) - L(x, y, \sigma)\end{matrix} \qquad (3)$$

The next stage is to perform a detailed fit to the nearby data for location, edge response and peak magnitude. A location in image scale space is identified that are invariant with respect to image rotation, translation and scaling. At each candidate location, a detailed model is fit to determine location, scale and contrast. Keypoints are selected based on measures of their stability.

2.Orientation assignment:

One or more orientations are assigned to each keypointlocation based on local image properties. All future operations are performed relative to the assigned orientation, scale and location for each feature, providing invariance to these transformations.

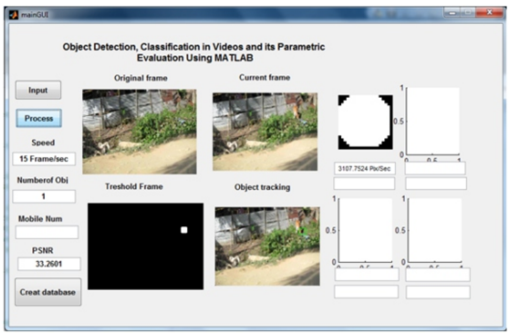
3.Keypoint descriptor:

The local image gradients are measured at the selected scale in the region around each keypoint and transformed into a representation that allows for local shape distortion and change in illumination. Figure 2 illustrates the computation of the keypoint descriptor as described in [10]. A Gaussian weighting function with σ equal to one half the width of the feature-vector window is used to assign a weight to the magnitude of each sample point.

VI. Results and Analysis:

This method was tested on several video sequences. All images were of size 324×244 with dynamic background. The proposed system was able to detect and human identification in almost all of the moving objects in the sequences. And it is implemented with the help of MATLAB tool and run on desktop PC with a 3.09GHz Intel and 1.90 GB RAM. Figure shows a sample input video is uploaded with dynamic background and the frame conversion is applied and finally the output is obtained and parameters are evaluated.

In matlab we have implemented this image subtraction, threshold and pattern matching algorithm. In the figure , the man is moving behind the tree it will detect and track movement finally classified as human or not. Then we calculated speed and velocity of object using matlab command and then found out the distance according to reference point. The detection output is more accurate and thus can significantly improve the performance of video surveillance based applications.



(a)

```
message =

    'Object enterd_Cow'

sending message
```

(b)

Fig. 3 Sample Output Result. (a) GUI view for Object detection and velocity calculation (b) Object name in command window

VII. Conclusion and Further Work:

This paper implemented a novel efficient algorithm for object detection and classification in video for surveillance applications. The strength of our approach lies in the ability to separate background and foreground in accurate and implemented in simple way with low time consumption and less noise, and proved effectively in both static and dynamic background texture scenes. The velocity and speed of the moving object can also be calculated. In future work, we will incorporate the complex video like disaster videos for object detection and classification present in the scene for more sophisticated vision based applications like fire accident, earthquake disaster etc.

References:

1.Xiaowei Zhou, Can Yang, and Weichuan Yu, "Moving Object Detection by Detecting Contiguous Outliers in the Low-Rank Representation," IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. 35, No. 3, March 2013.

2.Fei Chen, Huimin Yu, and Roland Hu, "Shape Sparse Representation for Joint Object Classification and Segmentation," IEEE Transactions On Image Processing, Vol. 22, No. 3, March 2013.

3.F. Lecumberry, A. Pardo, and G. Sapiro, "Simultaneous object classification and segmentation with high-order multiple shape models," IEEE Trans. Image Process., vol. 19, no. 3, pp. 625–635, Mar. 2010.

4.Lijing Zhang, Yingli Liang "Motion human detection based on background subtraction," IEEE Trans. Second International Workshop on Education Technology and Computer Science, March 2010.

5.J. Mike McHugh, JanuszKonrad, VenkateshSaligrama, and Pierre-Marc Jodoin, "Foreground-Adaptive Background Subtraction," IEEE Signal Processing Letters, Vol. 16, No. 5, May 2009.

6.BinengZhong, Shaohui Liu, Hongxun Yao, Baochang Zhang, "Multi-Resolution Background Subtraction For Dynamic Scenes," IEEE Trans., Nov. 2009.

7.Daniel Cremers, Stanley J. Osher, Stefano Soatto, "Kernel Density Estimation and Intrinsic Alignment for Shape Priors in Level Set Segmentation," International Journal of Computer Vision 69(3), 335–351, 2006.

8.A. Mittal and N. Paragios, "Motion-Based Background Subtraction Using Adaptive Kernel Density Estimation," Proc. IEEE Conf. Computer Vision and Pattern Recognition, June 2004.

9.F. Ranchin, and F. Dibos, "Moving objects segmentation using optical flow estimation", in Proc. Workshop on Mathematics and Image Analysis, Paris, France, Sep. 2004.

10D. G. Lowe, "Distinctive image features from scale-invariantkeypoints," in International Jounal on Computer Vision, 2004, vol. 60, no. 2, pp. 91–110.

11.BirBhanu and Jing Peng, "Adaptive Integrated Image Segmentation and Object Recognition," IEEE Transactions On Systems, Man, And Cybernetics—Part C: Applications And Reviews, Vol. 30, No. 4, November 2000.

12.Chris Stauffer, W.E.L Grimson, "Adaptive background mixture models for real-time tracking," IEEE Trans on Computer Vision, 1999.

13.Ahmed S. Abutaleb, "Automatic Thresholding of Gray-Level Pictures Using Two-Dimensional Entropy," Computer Vision, Graphics, And Image Processing 47, 22-32 (1989).

14.Milan Sonaka, Vaclav Hlavac, Roger Boyle "Image Processing, Analysis, and Machine Vision" Second Edition, PWS Publishing.

15.R. C. Gonzalez University of Tennessee, Richard E. Woods MedData Interactive, "Digital Image Processing," Second Edition, Prentice Hall.