Automated Recognition Of Object In Digital Images And Real Time Videos

Dr.G.Nagappan
Professor
Department of Computer Science
Saveetha Engineering College
Nagappan.cse@saveetha.ac.in

Preethi K
Post graduate
Department of computer science
Saveetha Engineering College
preethukeshav@gmail.com

Abstract— **Object detection has seen huge process in recent years. Object detection is the process to determine whether there are instances of interest in an image. The main difficulty of object detection arises from high variability in appearance among objects of the same class. An automatic object detection system. must also to determine the presence and absence of objects with different sizes and viewpoints under complex background clutters. An object detection approach using histogram features is proposed. The Histogram of Oriented Gradient (HOG) is a feature descriptor used in computer vision and image processing. The influence of each stage of the computation on performance, concluding that fine scale gradients, fine orientation binning, relatively coarse binning and high quality local contrast normalization in overlapping descriptor blocks are all important for good results .Experimental results show that the proposed approach is for effective object detection has better accuracy with less processing time consumption rather than existing methods.**

Keywords: Object detection, histogram of oriented gradient, feature extraction, HOG, Support Vector Machine.

Introduction

Observing and reasoning about object motion patterns are key tasks in numerous real world application domains such as visual surveillance, robotics and intelligent vehicles. In fixed camera video surveillance, this could aid human operators to monitor many video streams and focus their attention on possible incidents. Among the vast variety of existing approaches to object recognition there is a remarkable success of methods using histogram based image descriptors. Detecting objects in images is a challenging task owing to their variable appearance and the wide range of poses that they can adopt. The first need is robust feature set that allows the human form to be discriminated cleanly, even in cluttered backgrounds under difficult illumination. The issue of feature sets for object detection, showing that locally normalized Histogram of Oriented Gradient (HOG) descriptors provide excellent performance relative to other existing feature sets including wavelets is studied. The HOG representation has several advantages. It captures edge or gradient structure that is very characteristic of local shape and it does so in a local representation with an easily controllable degree of invariance to local geometric and photometric transformations translations or rotations make little difference if they are much smaller that the orientation bin size. The contexts from the high level task perspective are considered. It has been demonstrated that the object detection and classification tasks can provide natural comprehensive context for each other without any external assistance, and thus can be mutually contextualized for performance boosting. It is intuitively straightforward that for object classification task, the information from the local appearance promotes the performance significantly. For object detection task, the global context from object classification helps the detector better eliminate the false alarm. Although there are some works focusing on this direction, the underlying improvements brought by the context models for both two tasks have been underestimation are noticed. And the previous works take the context model in a multi feature fusion fashion without dedicate d design. In this wok, developed a histogram scheme for object detection and classification based on the Support Vector Machine(SVM) method.

I. RELATED WORK

Due to its pervasiveness in various contexts, background subtraction has been afforded by several researchers and plenty of literature has been published. The usual approach to moving object detection is through background subtraction that consists in maintaining an up-to-date model of the background and detecting moving objects as those that deviate from such a model. Compared to other approaches, such as optical flow, this approach is computationally affordable for real-time applications. The main problem is its sensitivity to dynamic scene changes and the consequent need for the background model adaptation via background maintenance. Such problem is known to be significant and difficult. Some of the well known issues in background maintenance that will be specifically addressed in the sequel.

[1] Generalized Stauffer-Grimson background subtraction for dynamic scenes: An adaptive model for backgrounds containing significant stochastic motion is proposed.The new model is based on a generalization of the Stauffer-Grimson background model, where each mixture component is modeled as a dynamic texture. An online K-means algorithm for updating the parameters using a set test of sufficient statistics of the model is derived. Dynamic textures, Background models, Background subtraction, mixture models, adaptive models techniques are used. Drawbacks are only suitable for static scenes and there is a strong limitation for scenes with spatiotemporal dynamics.

[2] Tracking and Counting people in visual surveillance systems:

The greatest challenge on monitoring characters from a monocular video scene is to track targets under occlusion conditions. In this work, a scheme to automatically track and count people in a surveillance system is presented. First, a dynamic background subtraction module is employed to model light variation and then to determine pedestrian objects from a static scene. To identify foreground regions are treated as decision features. Moreover, the performance to track individuals is improved by using the modified overlap tracker, which investigates the centric distance between neighboring objects to help on target tracking in occlusion states of merging and splitting. On the experiments of tracking and counting people in three video sequences, the results exhibit that the proposed scheme can improve the averaged detection ration about 10% as compared to the conventional work. Intelligent surveillance system, people tracking, people counting , overlap tracker, occlusion techniques are used. Drawbacks are sensitive to light variations.

[3] A people counting system based on BP neural network

A people counting system based on a back propagation (BP) neural network is proposed in this paper. The proposed system uses cheap photoelectric sensor to collect data and introduces BP neural network for counting and recognition , it is effective and flexible for the purpose of performing people counting. In this paper, new methods for segmentation and feature extractor are developed to enhance the classification performance. Promising results were obtained and the analysis indicates that the proposed system based on BP neural network provides good results with low false rate and it is effective for people counting. Techniques are Segmentation, feature extraction and Back Propagation Neural Network. Drawbacks are to fully implement a standard neural network architecture would require lots of computational resources and they are black-box.

[4] Fuzzy color histogram and its use in color image retrieval

A Conventional Color Histogram (CCH) considers neither the color similarity across different bins nor the color dissimilarity in the same bin. Therefore, it is sensitive to noisy interference such as illumination changes and quantization errors. Furthermore, CCHs large dimension of histogram bins require large computation on histogram comparison. To address these concerns, this paper presents a new color histogram representation called fuzzy color histogram (FCH), by considering the color similarity of each pixel's color associated to all the histogram bins through fuzzy set membership function. A novel and fast approach for computing the membership values based on fuzzy means algorithm is introduced. The proposed FCH is further exploited in the application of image indexing and retrieval. Experimental results clearly show that FCH is yields better retrieval results than CCH. Such computing methodology is fairly desirable for image retrieval over large image databases. Here techniques used are conventional color histogram, illumination changes, Fuzzy-means, Fuzzy color histogram, Image indexing and retrieval, membership matrix. Drawbacks are representation is dependent of the color of the object being studied, Ignoring its shape and texture, High sensitivity to noisy interference such as lighting intensity changes and quantization errors.

[5] Tracking of entities in Multi-Sport Field

This paper describes a method to track object in a multi-sport field. Mean shift algorithm is a kernel tracking approach which tracks on the basis of color and shape of the target object. Due to constant motion in a playground, it is difficult to track the same body because of changing color and direction. It enables to calculate the distance covered by an entity in the field. Proposed to overcome the drawbacks of mean shift this does not remain effective when a collision between players occur. Drawbacks are the cooler intensities and the size of the object frame varies due to the random motion of the player.

II. PROPOSED WORK

Fig.1 gives the block diagram of the proposed system for object segmentation, extraction and detection. The main goal of its tracking object in unconstrained and cluttered environments as they form groups at low frame rate, interact and part from the group in presence of occultation. The proposed algorithm contains two steps: color segmentation and extraction. At color segmentation, to simplify the problem, assuming that the camera is stationary and the background model is static. The first step of the algorithm is histogram equalizer, which is performed on a block level. For high frame rate sequences, the adjacent frame histogram method is used as the motion change between consecutive frames is small. This method eliminates the stationary background, leaving only the desired motion regions.

III. IMPLEMENTATION

Divide the image into small connected regions called cells, and for each cell compute a histogram of gradient directions or edge orientations for the pixels within the cell. Discredited each cell into angular bins according to the gradient orientation. Each cell’s pixel contributes weighted gradient to its corresponding angular bin. Groups of adjacent cells are considered as spatial regions called blocks. The grouping of cells into a block is the basis for grouping and normalization of histograms. Normalized group of histograms represents the block histogram. The set of these block histograms represents the descriptor.

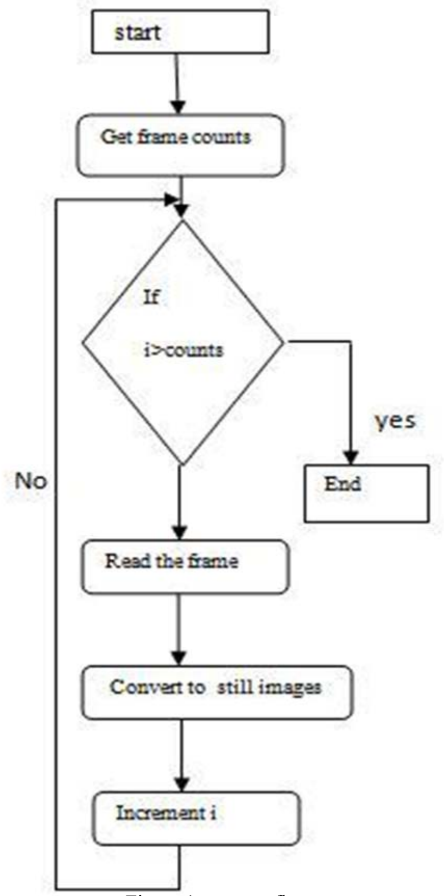


Figure 1: process flow

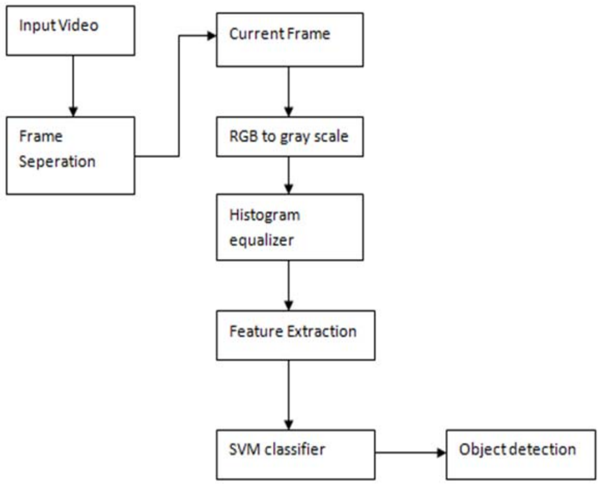


Figure 2: Block diagram of object detection

FRAME SEPERATION:

An input video(.avi files) is converted into still images for processing it and to detect the moving objects. These sequences of images gathered from video files by finding the information about it through 'aviinfo' command. These frames are converted into images with help of the command 'frame2im'.Create the name to each images and this process will be continued for all the video frames. In the Figure 1 diagram represents the process flow of this separation.

PREPROCESSING:

When the background model has been initialized through the training period the classification of pixels as foreground or background is a relatively simple task. If a pixel at image coordinate (x,y) lies within the subspace of RGB space described by any of the code words in the codebook at position(x,y), then the pixel is classified as background, otherwise it I foreground.

In the classification the ability to handle shadows and highlights since the model accounts for the intensity variation. The background model's way of dealing with chromaticity differences also gives the foreground/background classification, a high detection sensitivity to foreground objects that are similar in color to the background. In the classification process the values of min and max can be different from the values used during training. It is reasonable to only allow small variation in the chromaticities during the relatively short training period, whereas the variation over longer periods d be greater due to large changes in illumination conditions. This circumstance motivates the use small values of min and max during training and larger values during foreground/background.

The preprocessing , mentions that smoothing of the images can be used to reduce camera noise and remove transient environmental noise such as rain. Many algorithms use a Gaussian blur first to average out fluctuating pixel values to alleviate big differences. Alternatively, when temporal data can be exploited in a video, if a pixel's value is constantly changing over time then it can be assumed it is part of a non-static background object. The background model can deal with events such as objects changing positions by implementation an effective update rule to change the model over time.

HISTOGRAM EQUALIZER:

The next step in computing the descriptors is to normalize the histograms. On gradient vectors, can add or subtract a fixed amount of brightness to every pixel in the image, and as a result same gradient vectors at every pixel is found. It turns out that by normalizing gradient vectors, can make them invariant to multiplications of the pixel values. Dividing a vector by its magnitude is referred to as normalizing the vector to unit length, because the resulting vector has a magnitude of 1. Normalizing a vector does not affect its orientation, only the magnitude.

FEATURE EXTRACTION:

A good feature makes the classifier's job as easy as possible. It removes the irrelevant details to the classification. Local object appearance and shape can often be characterized rather well by the distribution of local intensity gradients or edge directions. It is implemented by dividing the image window into small spatial regions , for each cell accumulating a local 1-D histogram of gradient directions or edge orientations over the pixel of the cell. The combined histogram entries form the representation for better invariance to illumination and also useful to contrast-normalize the local responses before. This can also done by a accumulating a measure of local histogram "energy" over somewhat larger spatial regions("blocks") and using the results to normalize all of the cells in the block. It is referred to the normalized descriptor blocks as Histogram of Oriented Gradient (HOG) descriptors.

SVM CLASSIFIER:

Support Vector Machine(SVM) training algorithm builds a model that assigns new examples to one category to another making it a non-probabilistic binary linear classifier. An SVM model is a representation of the examples as points in space , mapped so that the examples of the separate categories are divided by a clear gap as wide as possible. It is classification task involves with training set contains one target values and several attributes. The goal is to produce a model which predicts target value of data instances in the testing set which are given only the attributes.

SVM Classification:

$$\mathbf{min}_{\mathrm{f_i}\boldsymbol{\varepsilon}_\mathrm{i}} \, ||\boldsymbol{f}||_{\boldsymbol{k}}^{\mathbf{2}} \quad + \boldsymbol{C}\sum_{\boldsymbol{i}=\mathbf{1}}^{\boldsymbol{l}} \boldsymbol{\varepsilon_i} \qquad y_i f \ \ (x_i) \geq 1 - \ \varepsilon_i \text{ for all i, } \varepsilon_i \geq 0$$

SVM classification, dual formulation:

$$\boldsymbol{min}_{\boldsymbol{\alpha_i}} \sum_{\boldsymbol{i}=\mathbf{1}}^{\boldsymbol{l}} \boldsymbol{\alpha_i} - \frac{\mathbf{1}}{\mathbf{2}} \sum_{\boldsymbol{i}=\mathbf{1}}^{\boldsymbol{l}} \quad \sum_{\boldsymbol{j}=\mathbf{1}}^{\boldsymbol{l}} \boldsymbol{\alpha_i} \, \boldsymbol{\alpha_j} \boldsymbol{y_i} \boldsymbol{y_j} \boldsymbol{K}(\boldsymbol{x_i} \boldsymbol{x_j})$$

$$\sum_{\boldsymbol{i}=\mathbf{1}}^{\boldsymbol{l}} \boldsymbol{\alpha_i} \, \boldsymbol{y_i} = \mathbf{0} \quad \mathbf{0} \leq \boldsymbol{\alpha_i} \leq \boldsymbol{c} \text{ for all i ;}$$

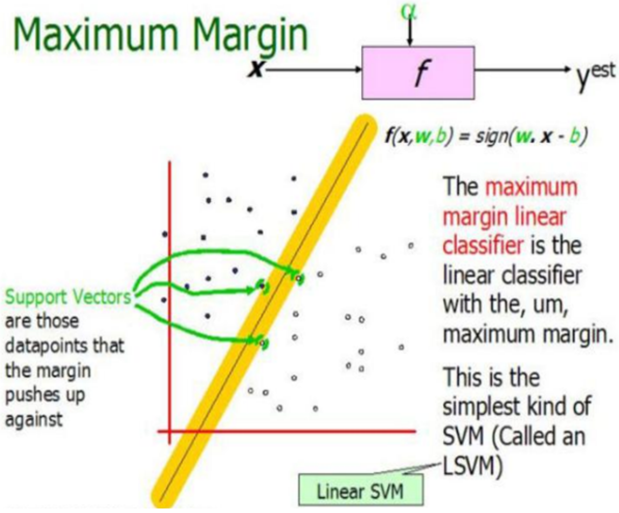
Variables $\boldsymbol{\varepsilon_i}$ are called slack variables and they measure the error made at point $(\boldsymbol{x_i}, \boldsymbol{y_i})$**.** Training SVM becomes quite challenging when the number of training points is large. For calculating the SVM, the goal is to correctly classify all the data. For Mathematical calculations,

[a] If $y_i = +1$; $wx_i + b \geq 1$

[b] If $y_i = -1$; $wx_i + b \leq 1$

[c] For all i ; $y_i(w_i + b) \geq 1$

$$\text{margin} \equiv \arg\min_{\mathbf{x}\in D} d(\mathbf{x}) = \arg\min_{\mathbf{x}\in D} \frac{|\mathbf{x}\cdot\mathbf{w}+b|}{\sqrt{\sum_{i=1}^{d} w_i^2}}$$



Linear SVM

IV. CONCLUSION

Traditional classification approaches perform poorly when working directly because of the high dimensionality of the data, but support vector machines can avoid the pitfalls of very high dimensional representations. The fact that SVM can perform as well as these systems without including any detailed prior knowledge is certainly remarkable. The influence of the various descriptor parameters and conclude that using support vector machine as a classifier fast detection with an excellent detection rate is built.

V.REFERENCES

[1] C. Stauffer , W.E.L. Grimson, “Adaptive background mixture models for real-time tracking,” In Proc. Computer Vision and Pattern Recognition conference ,pp. 246-252, 1999.

[2] W. Hu, T. Tan, “A Survey on Visual Surveillance of Object Motion and Behaviors”, IEEE Trans. Systems, Man, and Cybernetics, Vol. 34, No. 3, pp. 334-352, 2006

[3] Ju Han and Kai- Kuang Ma,” Fuzzy Color Histogram and Its Use in Color Image Retrieval,” IEEE transactions on image processing, Vol.11, NO.8 , August 2002

[4] Na-Na Li, Jie Song, Rui-Yui Zhou and Jun-Hua Gu,” A Person Counting System using BP Neural Network,” IEEE Fourth International Conference On Fuzzy systems and knowledge Discovery, 2007.

[5] Khyati K.Vyas , Kevin Vira and Ruhina Karani, “ Tracking of Entities in sport field, “ Elsevier International Conference on advanced computing technologies and applications,2015.

[6] S.C. Cheung and C.Kamath,” Robust techniques for background subtraction in urban traffic video, “video communications and Image processing , SPIE Electronic Imaging, UCRLConf.SanJose, Vol.200706, Jan 2004.

[7] Z. Zivkovic and F. van der Heijden,” Efficient adaptive density estimation per image pixel for the task of background subtraction”, Pattern Recognition Letters, Vol. 27 No.7, pp. 773–780, 2006.

[8] F-H. Cheng, Y-L. Chen, “Real time multiple objects tracking and identification based on discrete wavelet transform”, Elsevier journal of Pattern Recognition, Vol. 39, No. 6, pp. 1126–1139, 2006.

[9] A. Khare, U. S. Tiwary, W. Pedrycz, M. Jeon, “Multilevel adaptive thresholding and shrinkage technique for denoising using Daubechies complex wavelet transform”, The Imaging Science Journal, Vol. 58, No. 6, Dec. 2010, pp. 340-358.