ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 10, March 2016

SVM NOISE MARGIN OPTIMIZATION ALGORITHM FOR VIDEO BACKGROUND SUBTRACTION

*Stenwills**Frank***Xiao Fu

Senior Researcher, Linyi Top Network Co. Ltd, China ,
Junior Researcher, Linyi Top Network Co. Ltd, China
Juniorss Researcher, Linyi Top Network Co. Ltd, China

Abstract— **The important feature of detecting the moving objects in videos is Background subtraction. The main process involved in the background is the foreground detection. However, many algorithms usually neglect the fact that the background images consist of different objects whose conditions may change frequently. In this paper, a hierarchical background model is proposed based on segmenting the background images. It first segments the background images into several regions by the Support Vector Machine. Then, a hierarchical model is built with the region models and pixel models. The region model is extracted from the histogram of a specific region which is similar to the kind of a Gaussian mixture model. The pixel model is described by histograms of oriented gradients of pixels in each region based on the cooccurrence of image variations. We propose Silhouette detection algorithm. The experimental results are carried out with a video database to demonstrate the effectiveness, which is applied to both static and dynamic scenes by comparing it with some well-known background subtraction methods and according to the experiment, the Silhouette detection method is easy to operate and possesses high rate of accuracy, low rate of complexity, and well adapt to different kinds of shadow distribution.**

Keywords- **Background subtraction, hierarchical background model (HBM), pixel model, and region segmentation.**

I. INTRODUCTION

Background subtraction is a powerful and useful mechanism for detecting many changes in a sequence of images. There are many approaches for performing the background subtraction method. The main method involves the segmentation technique. [1]Background images are segmented into many regions and this segmentation is done by Mean-shift algorithm. From the segmented regions, we construct a hierarchical model that consists of region model and pixel model. The region model is mainly extracted from the histogram of a specific region which is similar to the Gaussian mixture model. The pixel model is based on the image of variation that occurs at the same time. These variations are described by the histogram of each region in pixels. As the images change frequently, the locations of background objects are not fixed, so each pixel of the segmented regions is assigned a weight to denote the probability that this pixel belongs to one region. The main disadvantage is that they usually neglect the fact that the background images consist of different objects whose conditions may change frequently.

The most widely used background subtraction methods are non-parametric and mixture-of-Gaussian models. The main task for designing a background subtraction algorithm which is fast is the way of selection done in a detection threshold. The threshold of varying video frames is selected by means of two models [2].There is usually a non-parametric model defined and in addition to it, a foreground model is introduced. There are two processes involved in Background subtraction that work in a loop: *background modelling* and *foreground detection*. In background modelling, the model in the view of a camera of the background is created and is periodically updated to handle the changes in illumination. In foreground detection, there involves two decision, first is made as to whether a new intensity fits the background model; second is the resulting *label field* which is fed back into background modelling to see that no foreground intensities is contaminating the background model. The inclusion of a foreground model tends to grow the detected regions rather than shrinking them. The main disadvantage of this paper is that the inclusion may lead to few false positives at the initial label field.

According to the functional characteristics of real-time image processing systems there were many existing digital image processing algorithms, but the most important Feature to have a reasonable hardware and software division for the realization of the functions existed [3]. With the basis of software and hardware division, a FPGA-based image processing system structure is built and respectively designs the structure of image acquisition and storage, image processing, real-time display and other functional modules. FPGA consists of four parts that involves lOB (input and output module), routing resources, logic unit, Block Ram Their main functions are:

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 10, March 2016

1) Input and output module: It is defined as the interface of the chip with the outside and is used to complete both the input and output of different electrical characteristics.
2) Logic unit: It is the main core used to complete all the logic functions.
3) Routing resource: It is used to connect the logic unit, lOB and Block Ram, mainly to achieve the good signal transmission.
4) Block Ram: It is used to achieve data storage.

Finally, the pre-processing circuit that is designed is experimentally verified, and the results show that the realization of the hardware design can meet the system functions and their time requiring for processing, which have certain practical value.

There is usually need of synchronization between the image acquisition and external trigger events. Hardware and software triggering are widely used although they have several limitations.[4] Soft synchronization is investigated in this paper to operate by time tagging both the trigger events and images in a video stream and thus by selecting the image corresponding to each trigger event. For the soft synchronization a stochastic model is developed and, based on the model, the uncertainty interval and confidence for correct image selection are determined, and an efficient and perfect calibration method is derived. For this method, the images are streamed from the camera to the image processing computer. This is where the images are time tagged, and then the image is selected which corresponds to the timing of an external trigger event. This method provides several advantages that include post-triggering capability and natural support for rapidly arriving trigger events, and increased flexibility with respect to how the trigger event is sensed and communicated, and reduced cost through the opportunity to use a data-network interface. The main disadvantage of soft synchronization is the timing resolution that is limited by the frame rate. First, the background images which are given as a input are segmented by mean shift segmentation [6]. Christo Ananth et al. [8] proposed a system in which the cross-diamond search algorithm employs two diamond search patterns (a large and small) and a halfway-stop technique. It finds small motion vectors with fewer search points than the DS algorithm while maintaining similar or even better search quality. The efficient Three Step Search (E3SS) algorithm requires less computation and performs better in terms of PSNR. Modified objected block-base vector search algorithm (MOBS) fully utilizes the correlations existing in motion vectors to reduce the computations. Fast Objected - Base Efficient (FOBE) Three Step Search algorithm combines E3SS and MOBS. By combining these two existing algorithms CDS and MOBS, a new algorithm is proposed with reduced computational complexity without degradation in quality.

Background is obtained by distribution. According to the experiments, susan method is easy to operate and possesses high rate of accuracy, low rate of complexity, and well adapt to different kinds of shadow distribution.

II. RELATED WORK

Background Subtraction is a process to detect a movement or significant differences inside of the video frame, when compared to a reference, and to remove all the non-significant components. A Hierarchical model is developed from the segmented regions of background using Mean-shift algorithm. This Hierarchical model consists of two models, region model and pixel model. The region model is mainly similar to that of the mixture of Gaussian and is mainly extracted from the histogram of regions which are specific. The pixel model is mainly made up of images that can take place at the same time and are connected to each other. The method proposed in this paper involves two processing levels. These are the steps taking place in existing system.

1. The frames of the video are segmented into regions by mean shift and are taken as the input.
2. Next, according to their position to form uniform segments for a scene region different frames are merged. When this procedure takes place, a dynamic strategy of representing region borders is also developed, which leads to a more robust performance for dynamic background.
3. Then the gray value histograms of these regions are computed to build the region models, and pixel models are computed by the pixel co occurrence within each region.
4. The region models are always built as Gaussian mixture models describing the number of components that are determined by a cluster algorithm.
5. For detecting foreground objects, we first usually segment an input frame according to the uniform segments determined.
6. Next, each region is detected whether it contains foreground objects by a corresponding region model.
7. If the detected result shows that a region contains foreground objects, first we will detect the pixel belonging to the foreground with the help of pixel models .Secondly after detecting each frame, parameters of region models and pixel models will be updated.

Advantages:

1. It’s not necessary that each model must be set constant parameters because assigning different parameters according to the region also leads to a more accurate description;
2. The weighted pixels in each region makes both the descriptor of region and pixels more precise; and
3. The hierarchical models reduces the time cost by just deciding which region contains the foreground and can avoiding other regions that doesn’t contain

Because in some dynamic scenes, the locations of background objects are not fixed, each pixel of the segmented

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 10, March 2016

regions is assigned a weight to denote the probability that this pixel belongs to one region.

Disadvantages:

- The main disadvantage is Noise. Mean-shift algorithm does not remove the complete noise in the background subtraction.
- The second is the shadows. Even the shadows are detected as moving object in the existing system.

III. PROPOSED SOLUTION

To overcome the problems faced in existing system, we use a advanced technique called Support-Vector Machine for segmenting the background images. The segmented images are meant to form a hierarchical model. This hierarchical model contains region model and pixel model.

The support vector machine is mainly based on statistic learning. .They are defined as supervised learning models with associated learning algorithms that can analyze the data and recognize the patterns, and also used for classification and regression analysis. It is a new machine learning theory. The support vector machine has been widely applied to many applications like pattern recognition, function approximation and system identification as because support vector machine is able to deal with both the classification and clustering

Classification and Regression by SVM

Generally the model can be divided into Support Vector Regression and Support Vector Classification. The Training data set is given as $\{(c_i,d_i)\}^N_{i=1}$ where ci R^n and corresponding binary class label b_i where a_i is the ith input vector with known binary target b_i, Let be a non-linear mapping from the data which are original to a high-dimensional feature space, and it is mainly used to replace sample points c_i , c_j and they have their mapping images as (c_i) and (c_j) respectively.

The weight and bias of hyper plane is defined as w and b, respectively. We define the hyper plane which may be ready to act as a decision surface in feature space, as such,

$$\Sigma^{i}_{i=1} w_i \phi_j (c) + b = 0 \qquad (1)$$

First we need to separate the data linearly in the feature space and so the decision function must meet a constraint conditions. The optimization problems are

$$Minimize\ \phi(\omega,\varepsilon) = 1/2\,\|w\|^2 + c\ \Sigma^{l}_{i=1} \varepsilon_i$$
$$Subject\ to\ d_i[(w.c_i) + b\ ] >= 1 - \varepsilon_i \qquad (2)$$

Where $_i$ is defined as a stack variable mainly used to relax the margin constraints which are hard. The regular constant C is mainly used to implement the trade-off between the maximal margin of separation and the classification error.

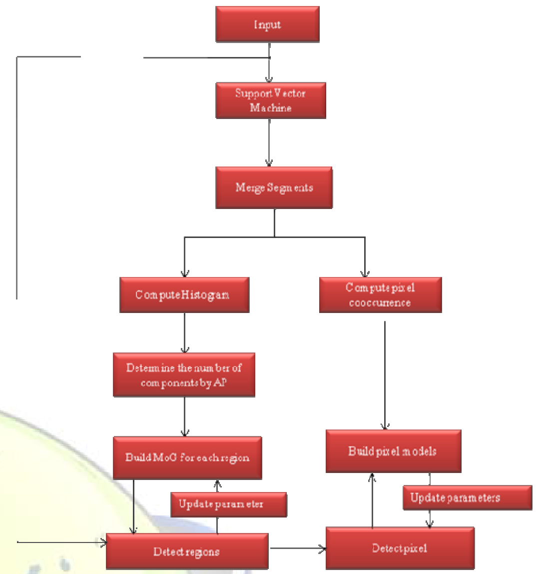

Figure 1: Block Diagram

Normally there is a problem of which data cannot be linearly separated in classification. So to avoid this problem, the Support Vector Machine can map the datas which are given as input into a feature space which are high deimensional.Usually the Support Vector Machine constructs an optimal hyper plane in the high dimensional space which is transferred into a non- linear decision frontier by first converting it into original space. The non-linear expression for the classification function is given as equation. (3)

$$t(x) = \Sigma^{l}_{i=1} a_i d_i\ K(c_i\,.\ c) + b \qquad (3)$$

The performance of SVM also includes the choice of this non-linear mapping function. The SVM applied uses the basis function to perform the operation of mapping. This function is expressed in (4).

$$K(c,d) = \exp(-s(c-d)^2) \qquad (4)$$

The s parameter in the above equation shows the reflection of the degree of generalization that is made to apply to the data used. Less generalization can be achieved in Support Vector Machine by obtaining more datas. When there is little s, it may reflect more generalization and a big one reflects less generalization. When the input data is not normalized, this parameter can perform a normalization task.

The classification scheme may be also defined with the case of the regression. In this case, the main idea for training the SVM by using d values different from +1 and -1. .Then, an approximation function is derived that fits approximately the known values only.

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 10, March 2016

Silhouette Detection Algorithm

Silhouette detection Algorithm is mainly used to detect the angle point of the image that is moving. This Algorithm is mainly involved to do the geometry observation mainly on the basis of image's gray-scale. Then it divides the pixels into 3 main points. They are angle point, edge point and flat area. To satisfy a different value when measured in different directions we need to apply round template to image.

The centre pixel of the template is always called as nucleus. While detecting the edge, we need to move the template which is in round to the image, and then compare every gray value of the pixel in the template along with the nucleus. If the D-value is smaller than the threshold value, mark that this point has similar gray-scale to the nucleus.

$$q(w,w_0)= \begin{cases} 1, & |O(r)-O(r0)|<= d \\ 0, & \text{else} \end{cases} \quad (5)$$

The parameter d is defined as the threshold value. If the d value is smaller than the threshold value, then that point is marked as to that similar gray-scale to the nucleus. q(w,w0) is defined as the pixel function. O(r0) is the gray-scale value of the nucleus that is in the center of template. O(r) is defined as the gray-scale value of other pixels in the templates. D is defined as the threshold value.

Therefore, for any image area the template goes through, the area which is formed by all the pixels to satisfy the formula (1) called as similar nucleus value area (USAN) [6]. The size of USAN area is as follows,

$$n\ (r_0)= \Sigma q(w\ ,\ w_0\ ) \quad (6)$$

There are two main aspects to consider by using Silhouette detection Algorithm for detecting image edge:

1)The template selection
2) To determine the value for d & g's threshold value.

The two elements used here are used to determine the efficiency of edges that are detected.

The template selection:
As the image is digitalized, the template found cannot be the real round .Hence to overcome this we use rectangle template (2m+1) * (2m+1) instead.

To determine the value for d & g's threshold value:
Threshold value D is used to determine the contrast ratio of the object and background that are recognizable. Area with smaller contrast ratio, D should be smaller area

IMPLEMENTATION AND EXPERIMENTAL RESULTS

This solution was implemented using MATLAB. It is a high-level language and interactive environment for numerical computation, visualization, and programming. Using MATLAB, you can analyze data, develop algorithms, and create models and applications. The language, tools, and built-in math functions enable you to explore multiple approaches and reach a solution faster than with spreadsheets or traditional programming languages. [11]

Performance Evaluation

It is important that after the application by the proposed model, has good results over existing techniques. The proposed method's performance is compared with many several available methods. The performance has been observed by calculating noise reduction in the background subtraction with 5 scenarios.

Scenario	Existing Algorithm	Our Algorithm
Moved Object	88.1	88.25
Time of Day	85.6	85.71
Traffic Area	78.5	78.54
Waving Trees	90.08	90.56
Foreground Aperture	93.4	93.56

Table 1: Experimental Result

The Experimental result shows that our proposed algorithm finds greater noise than the existing techniques.

IV. CONCLUSION AND FUTURE EXTENSIONS

We have presented a new background subtraction Technique just to extract foreground objects in a wide range of environmental conditions. Our new technique was designed to handle the problems typically associated with the background subtraction done by hierarchical model. The background subtraction done using Support Vector Machine solves the problem faced in the old technique. Our technique first uses statistical background subtraction to identify regions-of-interest containing the foreground object in the environment. Then a statistical model is formed from our technique. To further improve our results, we plan to include motion information into the saliency map, and employ shaped-based models for better figure completion and tracking. Through the experiment, when n value equals to 5*5, it will have the sound effect on the accuracy and efficiency.

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 10, March 2016

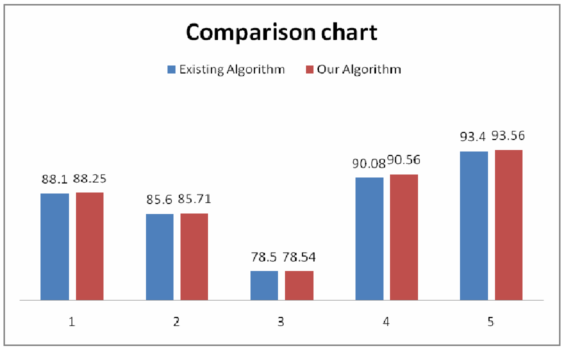

Figure 2: Comparison chart

V. ACKNOWLEDGMENTS

We wish to express our sincere thanks to all the staff member of C.S.E Department Loyola ICAM College of Engineering and Technology for their help and cooperation.

References:

[1] "A Hierarchical Model Incorporating Segmented Regions and Pixel Descriptors for Video Background Subtraction" Shengyong Chen, *Senior Member, IEEE*, Jianhua Zhang, *Member, IEEE*, Youfu Li, *Senior Member, IEEE*, and Jianwei Zhang, *Member, IEEE*

[2] "Foreground-Adaptive Background Subtraction" J. Mike McHugh, *Member, IEEE*, Janusz Konrad, *Fellow, IEEE*, Venkatesh Saligrama, *Senior Member, IEEE*, and Pierre-Marc Jodoin, *Member, IEEE*

[3] "FPGA-based Video Image Processing System Research" Zhao Wenge,He Huiming School of Information Engineering,Handan College

[4] "Soft Synchronization: Synchronization for Network Connected Machine Vision Systems" Brian S. R. Armstrong, *Senior Member, IEEE*, and Sandhya K. Puthan Veettil, *Member, IEEE*

[5] "Moving shadow detection based on Susan algorithm" Huang Si-ming Liu Bing-han College of Mathematics and Computer Science FuZhou University Fuzhou, China huangsu0703@sina.com Wang Wei-zhi College of Civil Engineering FuZhou University Fuzhou, China wwz@fzu.edu.cn

[6] MA Guizhen, FANG Zongliang, YAO Zongzhong. Performance Analysis and Comparison of SUSAN Edge Detector[J]. Modern Electronics Technique, 2007(8): 189-191

[7] R. Horst and M. Negin, "Vision system for high-resolution dimensional measurements and on-line SPC:Web process application," *IEEE Trans. Ind. Appl.*, vol. 28, no. 4, pp. 993–997, Jul.-Aug. 1992.

[8] Christo Ananth, A.Sujitha Nandhini, A.Subha Shree, S.V.Ramyaa, J.Princess, "Fobe Algorithm for Video Processing", International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering (IJAREEIE), Vol. 3, Issue 3,March 2014 , pp 7569-7574

[9] E. Kirda, C. Kruegel, G. Vigna, and N. Jovanovic. Noxes: A client-side solution for mitigating cross site scripting attacks. In SAC '06: Proceedings of the 2006 ACM symposium on Applied computing, pages 330–337, 2006. ACM.

[10] M.S.Lam,J.Whaley,V.B.Livshits,M.C.Martin, D. Avots, M. Carbin, and C. Unkel. Context sensitive program analysis as database queries. In Proceedings of the Twenty-fourth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems. ACM, June 2005.