ISSN XXXX XXXX © 2017 IJESC

Research Article Volume 7 Issue No.5

Anomaly Detection using CCTV in Closed Environment

Kaushal Savla[1], Priyank Shah[2], Rushabh Shah[3], Prof. Sai Deepthi Pabba[4]
BE Student[1, 2, 3], Assistant Professor [4]
Department of Computer Engineering
K. J. Somaiya College of Engineering, Mumbai, Maharashtra, India

Abstract:
As a result of the immediate access to money and client information, mechanized teller machines (ATMs) are the objective of complex assaults and misrepresentation. To counter this issue, present day ATMs use specific equipment security frameworks that are intended to identify specific sorts of assaults and control. Be that as it may, such frameworks do not give any insurance against future assaults that are obscure at configuration time. In this paper, we propose an approach that can identify referred to and obscure assaults on ATMs and that does not require extra security equipment. The thought is to use programmed demonstrate era methods to learn examples of ordinary conduct from the status data of standard gadgets involved in an ATM; a critical deviation from the educated conduct is a pointer of an extortion endeavor. We cast the distinguishing proof of ATM misrepresentation as a sequence based inconsistency location issue, and we depict three particular strategies that actualize our approach.

Keywords: Anomaly Detection, Object Detection, Movement Detection, Machine learning.

I. INTRODUCTION

Computerized teller machines, ATMs, are liable to different assaults. The essential purpose behind this is the sum of money inside the ATM safe (up to 5, 00,000 in a high volume ATM), additionally the entrance to client information that thusly offers access to money. The aggregate misfortunes from ATM extortion amid 2008 crosswise over Europe are assessed to 485.15 million EUR.1 This clarifies the counteractive action against assaults and extortion is a theme of most noteworthy significance, not just for budgetary establishments furthermore, bank clients additionally for ATM producers. Present day ATMs contain an assortment of security innovation. Inner gadgets either goes about as self-sufficient high security modules or is ensured by encryption also, the encompassing safe. Moreover, present day ATMs contain specific security sensors extending from shake sensors over gas sensors to cameras. The security condition of an ATM is typically checked by a product framework to recognize assaults and to respond as needs be. The distinguishing proof occurs in a knowledgebase way, i.e., in light of an arrangement of master decides that should be indicated physically for individual assaults. After particular security equipment and specific master guidelines can be produced after a novel assault has turned out to be known, updating all ATMs in administration is a period and cost-escalated handle. Besides, such contraventions must be started after the assault has happened, i.e., after misfortune or harm has as of now been caused. We propose a novel way to deal with recognize ATM misrepresentation that overcomes the said impediments of current security arrangements. Rather than demonstrating (known) assaults, we show the typical conduct of an ATM. This depends on the presumption that a critical deviation from the typical conduct is a solid marker of an assault (which require not really is known already). Also, rather than physically determining master rules, we handle the issue from an information driven perspective via naturally creating a model of ordinary conduct in light of the information stream of status data delivered by the equipment gadgets and the programming segments inside an ATM. Compared to current security solutions, our approach has the following benefits:

1. It does not require particular (security) hardware, but uses the prevailing hardware equipment.
2. It can be applied on any machine, independent of the type, the equipment, and the manufacturer.
3. It minimizes human effort

II. PROPOSED MODEL

The model that we proposed will be able to detect the anomaly event in a closed environment under various conditions.

1. Abnormal Behaviour (rapid motions)
2. Person with mask.
3. Harmful Object in ATM (gun, knife, etc.)

III. IMPLEMENTATION OF MODEL

The machine is learning in this step. We train the machine by giving the input of the initial environment. The machine now knows what an ideal environment is. Once the machine knows the normal environment of a system, we can now test it. In testing, we would be checking for three cases viz. Motion Detection, Skin Detection and Object Detection. The threshold values are set while initial training of environment which would be now compared with the real world input helping in detection of any anomaly if present. The threshold values are determined by initial testing the machine with various test cases. We take a snapshot of current frame from live video input, which will now be further processed. The current frame is compared with initial trained image. If the values are greater than a particular threshold then it would come to know that there is some motion in the system, whether normal or abnormal. Here we are calculating two different type of motion. The first is by calculating difference between the current frame and the initial frame with which the machine is trained. And the second one is the difference between current frame and the previous frame. The first difference is used to identify any type of motion in the environment which can be normal as well as abnormal. This is done because all the time there won't be any motion in the system. If this difference is greater than a particular threshold then we will be further proceed to calculate our next difference. This difference is to check whether there is abnormal event in the system. If this difference is greater than particular threshold, there would be an alarm denoting the occurrence of abnormal event in the system. If the detected motion is termed as normal by the machine then the machine will check for the percentage of the skin in the current frame. If it is less than a threshold value then there will be detection of abnormal event. Further it will also check for all the harmful objects which may be present in the system. A database of all harmful objects is maintained and fed to machine for matching.

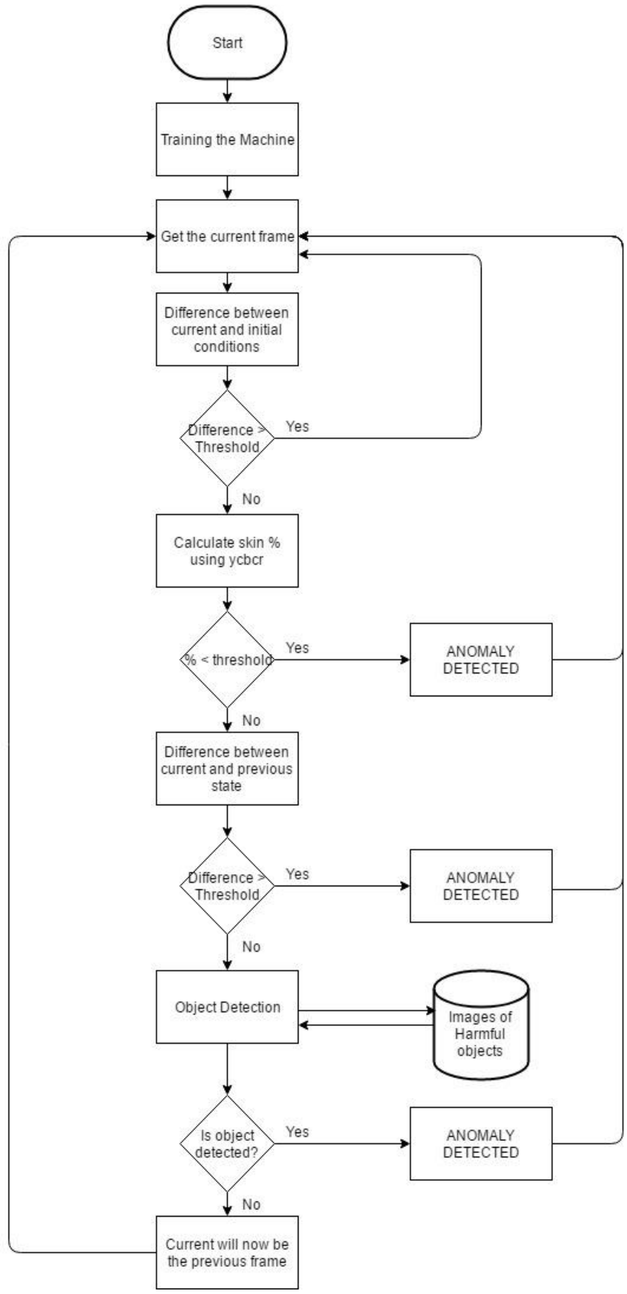

Figure.1. machine for matching

A. The flow is divided into four parts as follows:

A. Train the machine

The machine is learning in this step. We train the machine by giving the input of the initial environment. The machine now knows what an ideal environment is.

B. Motion Detection

Motion detection is one of the key strategies for programmed video examination to extricate vital data from scenes in video surveillance systems. Motion detection is one of essential assignment in video processing and scenes understanding frameworks. It separates pivotal data from scenes which is utilized as a part of numerous PC vision applications, for example, Automatic video surveillance, object tracking and classification, action understanding and so on. This makes motion detection an exceptionally dynamic research range in computer vision and its execution in automated visual surveillance systems... Visual observation is a key innovation for fight against terrorism, crime and public safety.

C. Skin Detection using YCbCr

YCbCr colour space has been defined to meet the increasing requirements of digital algorithms in handling video information and has become the most frequently used colour space in digital videos. YCbCr is consisting of three components, two of them are of chrominance and one is of luminance. So as to enhance/optimize the performance of skin colour clustering, the present work utilizes YCbCr space to assemble a skin colour model, since it is additionally referred to that, as the chrominance parts are practically free of luminance segment in the space there are non-straight relations between chrominance (Cb, Cr) and luminance(Y) of skin pixel shading in the high and low luminance locale. As in RGB space, the triple part (r,g,b) communicates shading as well as luminance. Luminance may change over a man's face because of the information and has become the most frequently used colour Space in digital videos. YCbCr is consisting of three components, two of them are of chrominance and one is of luminance. So as to enhance/optimize the performance of skin colour clustering, the present work utilizes YCbCr space to assemble a skin colour model, since it is additionally referred to that, as the chrominance parts are practically free of luminance segment in the space there are non-straight relations between chrominance (Cb, Cr) and luminance(Y) of skin pixel shading in the high and low luminance locale. As in RGB space, the triple part (r,g,b) communicates shading as well as luminance. Luminance may change over a man's face because of the surrounding lighting and is not a solid measure in separating skin from non-skin area. YCbCr is actually an encoded nonlinear RGB signal, commonly used by European television studios and for image Compression work. Colour is shown by luma (that is luminance, registered from nonlinear RGB), it is built as a weighted aggregate of the RGB qualities, and two colour distinction values Cr and Cb that are figured by subtracting luma from RGB red and blue segments. The transformation effortlessness and express detachment of luminance and chrominance segments makes this colour space appealing for skin colour modelling. In YCbCr colour space, the two Chroma segments Cr, and Cb can be productively used to characterize expressly skin locale. The limits be chosen as (Crmax; Crmin) and (Cbmax; Cbmin) , a pixel esteem is delegated skin pixel, if the qualities (Cr,Cb) fall inside the edges. The luminance must be expelled from the colour portrayal in the chromatic colour space. Chromatic hues known as "immaculate" hues without luminance. The YCbCr change from RGB shading space can be expert by taking after network.

$$\begin{bmatrix} Y \\ Cb \\ Cr \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 0.148 & -0.291 & 0.439 \\ 0.439 & -0.368 & -0.071 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} + \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix}$$

For the skin coloured pixel using YCbCr, red chrominance value lies between (140, 165), blue chrominance lies between (140, 195) and hue lies between (0.01, 0.1). The result using YCbCr as shown from results of YCbCr, it is clear that the segmented and colour segmented images are formed on the basis of skin coloured pixel. The further processing is also carried out. The skin region detected is shown in image.

Figure.2. C.1 (Left) input image, (Right) Skin detected

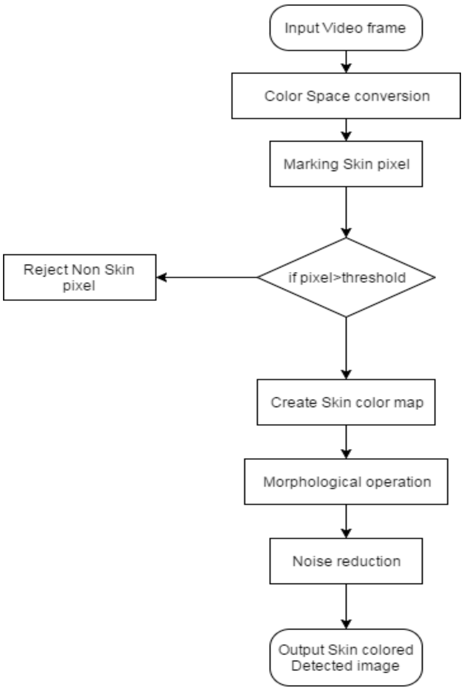

Figure.3.C.2 Flow for YCbCr

D. Object Detection using SURF

The Speeded-Up Robust Features (SURF) indicator descriptor calculation created by Bay et al is planned as a proficient other option to SIFT. It is speedier, and more robust when contrasted with SIFT. For the discovery of ideal Gaussian derivatives, this depends on basic 2D box channels; it utilizes a scale invariant blob locator in light of the determinant of Hessian framework for scale determination and areas. It's thought is to estimate the second order Gaussian subordinates in an effective path with the assistance of fundamental pictures utilizing an arrangement of box channels. The 9 × 9 confine channels delineated Fig 10 are approximations of a Gaussian with σ =1.2 and speak to the most minimal scale for registering the blob reaction maps. These approximations are meant by Dxx, Dyy, and Dxy. In this manner, the approximated determinant of Hessian can be communicated as

$$\det(H_{approx}) = D_{xx}D_{yy} - (wD_{xy})^2$$

Where w – (relative weight for the filter response) it is used to balance the expression for the Hessian's determinant. The approximated determinant of the Hessian speaks to the blob reaction in the picture. These reactions are put away in a blob reaction guide, and neighbourhood maxima are distinguished and refined utilizing quadratic addition, as with DoG. At long last, do non-most extreme concealment in a 3 × 3 × 3 neighbourhood to get interest points and the scale of values. The SURF descriptor begins by building a square district loped around the interest point, and situated along its fundamental introduction. The span of this window is 20s, where s is the scale at which the intrigue point is distinguished. At that point, the intrigue area is additionally isolated into littler 4 × 4 sub-locales and for each sub district the Harr wavelet reactions responses in the vertical and horizontal directions (meant dx and dy, separately) are figured at a 5 × 5 inspected focuses as appeared in Fig. 11

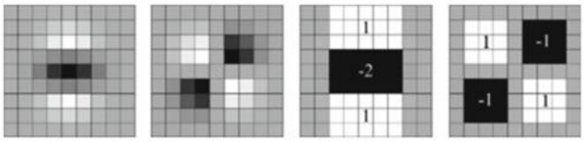

Figure.4.D.1 Left to right Gaussian second order derivatives in y- (Dyy), xy-direction (Dxy) and their approximations in the same directions, respectively.

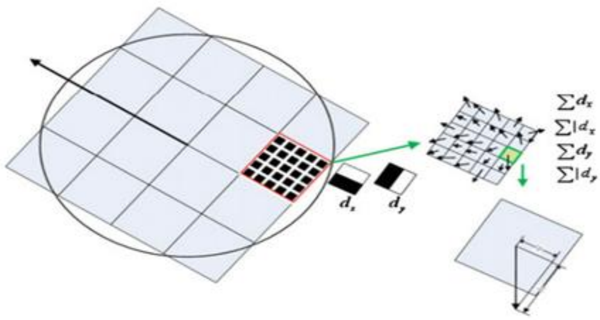

Figure.5.D.2 dividing the interest region into 4 × 4 sub-regions for computing the SURF descriptor

These responses are weighted with a Gaussian window centred at the interest point to increase the robustness against geometric deformations and localization errors. The wavelet responses dx and dy are summed up for each sub-region and entered in a feature vector v, where

$$v = \left(\sum d_x, \sum |d_x|, \sum d_y, \sum |d_y|\right)$$

Computing this for all the 4 × 4 sub-regions, resulting a feature descriptor of length 4 × 4 × 4 = 64 dimensions. Finally, the feature descriptor is normalized to a unit vector to reduce illumination effect. The primary reason for using SURF descriptor contrasted with SIFT is the handling speed as it uses 64 dimensional component vector to depict the local feature, while SIFT utilizes 128. Be that as it may, the SIFT descriptor is more appropriate to depict pictures images affected by translation, rotation, scaling, and other illumination deformations. In spite of the fact that SURF demonstrates its potential in an extensive variety of PC vision applications, it likewise has a few inadequacies. Whenever 2D or 3D items are thought about, it doesn't work if turn is fierce or the view edge is excessively unique

IV. EXAMPLE

If the motion of the user is greater from a given threshold then an abnormal event is detected. If the person wears mask or helmet i.e. if the skin percentage is lower than a threshold then an abnormal event is detected. If person brings harmful objects like knife or gun then an abnormal event is detected.

V. RESULTS AND DISCUSSION

Our ATM fraud detection approach depends on the presumption that a significant deviation from the ordinary conduct is a strong indicator of an assault. We use the information stream created inside an ATM to naturally produce a model of normal behavior, which is then used to recognize anomalies (or assaults separately).

VI. REFERENCES

[1] . M. Aly, P. Welinder, M. Munich, and P. Perona. Scaling Object recognition: Benchmark of current state of the art Techniques. TN First IEEE Workshop on Emergent Issues in Large Amounts of Visual Data (WS-LAVD), IEEE International Conference on Computer Vision (ICCV), September 2009. 6

[2] H. Bay, T. Tuytelaars, and L. Van Gool. Surf: Speeded up ro-bust features. In European Conference on Computer Vision, May 2006. I, 2

[3] M. Brown, S. Winder, and R. Szeliski. Multi-image matching using multi-scale oriented patches. In Computer Vision and Pattern Recognition, pages 510--517, 2005. 2

[4] D. Lowe. Towards a computational model for object recognition in IT cortex. Proc. Biologically Motivated Computer Vision, pages 2031, 2000.

[5] C. Evans. Open surf. http://www. chrisevansdev. com/ computer-visionpensurf.html, retrieved 11/04/2011.

[6] D. Chai, and K.N. Ngan, "Face segmentation using skin-color map in videophone applications". IEEE Trans. on Circuits and Systems for Video Technology, 9(4): 551-564, June 1999.

[7] M. J. Jones and J. M. Regh, “Statistical color models with application to skin detection”, In Proc. Of the CVPR ’99, vol. 1, 274-280, 1999.

[7] Kakumanu, Praveen, Sokratis Makrogiannis, and Nikolaos Bourbakis. "A survey of skin-color modeling and detection methods." Pattern recognition 40.3 (2007): 1106-1122.

[8] Prajkata M.Patil, Y.M.Patil, “Robust Skin color detection algorithm”. International journal of Engineering Research and technology(IJERT) Vol 1, issue 8, October 2012.

[9] Guoliang YangHuan Li, Li Zhang, Yue Cao ,” Research on a Skin Color Detection Algorithm Based on Self-adaptive Skin Color Model” International Conference on Communications and Intelligence Information Security, pp 266-270. October 2010. Space, RGB Color. "Color Spaces." (2013)

[10]http://www.compression.ru/download/articles/color_space/ch03.pdf.

[11] Yuh-Rau Wang, Wei-Hung Li, Ling Yang, ”A Novel real time Hand Detection based on skin color”, 17th IEEE International symposium on consumer electronics (ISCE), pp 141-142, 2013.