Abnormal Activity Detection and Alert System in ATM Centers

Alekhya M.K.[1] G. Niharika[2] M Prathibha[3] Joshi Priyanka[4] Dr Chidananda H[5]
[1,2,3,4,5]Department of Computer Science and Engineering
[1,2,3,4,5]Rao Bahadhur Y Mahabaleshwarappa Engineering College, Ballari, India

Abstract— Now-a-days Automated teller machines (ATM) are widely being used for banking transactions and are becoming one of the necessities of everyday life. ATMs facilities like withdrawal, deposit, and transfer of money from one account to another account have become a part of life. However, this convenience is ruined by criminal activities like money snatching and attack on customers, which are increasingly affecting the security of bank customers in daily life. In this paper, we propose a video based framework that efficiently identifies abnormal activities happening at the ATM centers and an alarm is generated during any unusual incidence. A Video frame Histogram difference is proposed method used to identify the abnormal activities from video sequence. The proposed framework is able to distinguish the normal and abnormal activities like money snatching, harm to the customer, or attack on the customer.

Keywords: abnormal activity, activity detection, alert system, ATM center

I. INTRODUCTION

The quick headway in Automatic Teller machines (ATM) has made life direct for the standard day by day presence, beside it isn't so for executives who oversee it. ATMs are not attested by banks rather they are outsourced to Managed service providers (MSP) from acquiring to keeping up the machines. A few sections like the assistance, cash filling, and security and along these lines the uninvolved resources inside the ATM rooms are responsible for keeping the ATM dynamic. Normally, an ATM site contains wherever between 8 to 12 uninvolved resources which join two air course and cooling structures, two light collecting sheets, Associate in Nursing inverter/UPS, a surveillance camera and no under eight to twelve lightweight globules. Security is one of the key worries of the edge time. The most significant issue of utilizing electronic gadgets and mechanical gathering is that the data can be hacked and spilled, security overlooked, robbery and theft. ATMs are general saving money associations which benefits especially precious for the clients. Most ATMs are open 24 hours; their zones are spread wherever on a city/town. It may be difficult to take cash from a bank, as it is equipped with high manual security and is by and large organized close populated regions of a city/town. Regardless, ATMs because of the likelihood of its association, is under a more basic danger of being burgled. Since the whole clarification behind an ATM advantage is to diminish manual connection and offer association to the client through robotization, it is fundamental that the security gave is in like way tweaked and beneficial. The tenacious reports of burglary and theft in ATM are a making worry for the banks. These security issues are changing into a hindrance to the development of ATM association and better resolute quality. The security structure should be helpfully robotized so it can perceive bizarre conditions inside the ATM and answer to the experts or jar the machine from discharging cash.

ATM is a computerized telecommunication device that serves the customer of a financial firm with a swift access to financial transactions in a public space by exempting the need for a clerk or bank teller. The numbers of ATM installations are increasing dramatically to support the transactions in billions. Increase in nefarious activities like robbery, murder, and other crimes have raised an urgency to install an effective system that can protect people as well as ATM installations [1, 2].

Generally ATM installations are equipped with CCTV cameras that keep a watchon the activities. Unfortunately, CCTV is not sufficient to provide security due to their inability to recognize unusual behaviours themselves [3] and hence monitoring authority needs to monitor these feeds 24 × 7 which is a challenging task. Today, we need an advanced system that can effectively monitor and automatically recognize unusual crime activities in an ATM room and can also report to the nearest monitoring firm before an offender could elope. Another approach to handle this situation could be an alarm system or electrical buzzer. Each ATM premise can be equipped with an electric buzzer. ATM user can press this buzzer to send signal to response group if any abnormal event takes place. Alarm systems may become ineffective as individual alarm must be responded by a main alarm response group, which should first examine the type or nature of the event being alarmed before any help signal can be requested. In addition, most alarms require a noticeable effort to operate, presenting an uncertainty that the perpetrator can simply physically stop the victim from triggering the alarm or may take a belligerent action against the victim if the victim is seen to initiate an alarm signal. Absence of automated security mechanism leads to post-incident forensic analysis by the law agencies.

(a) (b)

(c)

Fig. 1: Attack on person at various ATM installations.

Many a time law enforcement authorities become aware of the crime after several hours after the incident. This is a major problem in the urban areas as well as in the rural areas. Recently, a gruesome attack on a woman at an ATM located in Bangalore city, India [4], has brought to focus the issue of security at such kiosks (Figure 1(a)). This incident has sent shock waves across the country and highlighted the need to tackle such brutal acts. In some cases, ATM guard is also killed when he tries to save the victim because attackers are generally equipped with weapons like machete, guns, pistols, iron and rod and usually are multiple in numbers. Figure 1(b) [5] depicts the typical scene of a guard tied down with a rope by two attackers in Bangalore city, India, and Figure 1(c) [6] depicts the attack on a man at Karachi city, Pakistan. Therefore, it is necessary to have an automated system that can proactively identify and generate alarm on unusual behaviour.

Video based human activity analysis has gained lots of attention amongst the researchers. The goal of human activity recognition is to analyse different activities automatically from an unknown video [7]. Analysis of various activities involves recognition of motion pattern and generation of high level description of actions. There are various approaches like manifold approaches, spatiotemporal interest of feature points, motion history images, accumulated motion image, and bag of words model which are recently used by many researchers for effective human action recognition and representation [8–12]. In this paper, we present a system that can amend the current trends of the surveillance system. The system can automatically recognize different actions or number of persons through a CCTV camera like single normal, multiple normal, and multiple abnormal and generate signal accordingly. Using our system, the offender is more likely to be caught by the police red-handed because they are informed about the crime instantly. In addition, the proposed system can be used to generate automated alarm that can alert security guard deputed at the ATM location as well as other people around the premise to obtain immediate security. The paper is organized as follows. Section 2 presents related works and background of this work. In Section 3, we present our proposed method. Section 4 depicts the results and analysis of the proposed approach. Finally, conclusions are drawn in Section 5.

II. LITERATURE REVIEW

The intricacy at ATM booth described by COPS [1, 2] is the main motivation behind this research which has inspired us to develop an effective security system. In this section, we present the related work and research undergone in developing video based security system that helped us to make an efficient surveillance system. Various approaches have been proposed by researchers for human action recognition (HAR). Davis and Bobick [13] in their paper have presented the usage of temporal templates for recognizing human actions. References [7, 14] have presented a detailed survey on human motion and behavior analysis using MHI and its variants. Other approaches like Optical Flow and Random Sample Consensus (RANSAC) by [8] decipher the representation and recognition technique of human actions. For feature extraction, Hu has proposed a novel theory popularly known as Hu moments which are invariant to translation, scale, and rotation [15]. Bobick and Davis [16] in their paper have shown the usage of Hu moments for feature extraction from temporal templates. Hu moments are widely used shape descriptors due to its simplicity and less computational approach [17–19]. Various other descriptors like Fourier descriptors (FD) and Zernike moments have also been proposed. Fourier descriptors prove to be a disadvantage when the image size varies because the number of points also varies and the method becomes computationally high to work at real time. Zernike moments are advanced version of Hu moments whose magnitude is invariant to rotation but their computation time is also extensive to work at real time [20]. Besides the availability of various methods for feature extraction, we have used the conventional Hu moments method for shape description of the MHI/MEI. It is because Hu moments are computationally effective as compared to other descriptors. To make machine learn these features, a classifier has to be used. There are varieties of classifiers available like support vector machine (SVM), neural networks (NN), and Bayesian classifier. Debard et al. [21] have presented the identification of abnormal event, that is, fall using SVM. References [22–25] have shown a great adaptive learning of support vector machine in video surveillance. SVM, apart from its learning from two classes [26], has shown multiclass classification through SVM which helped us to analyze multiple classes through it. Sometimes it does happen that redundancy in data comes inherently from the video. For instance MHI/MEI formed by the presence of two persons in a video is also formed by the presence of an obese person. This kind of data may reduce the learning accuracy of SVM. So to address this kind of problem, principal component analysis (PCA) has been used. Reference [27] has shown the use of PCA with SVM in the work in action recognition in video. Another great work from [28, 29] has illustrated the use of PCA in dimension reduction. The main motive of this paper is to build a strong security framework system which can work at real time environment at ATM booth or other similar premises.

III. PROPOSED METHODOLOGY

The following block diagram shown below in fig 2 is the methodology of the proposed method:

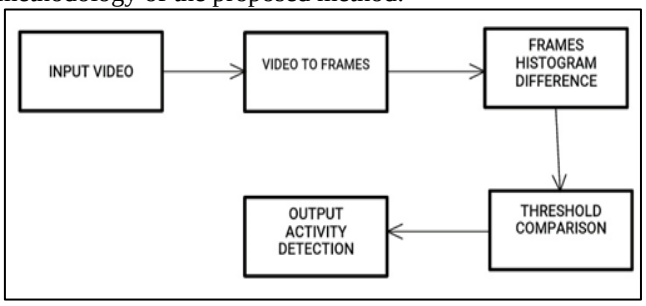

Fig. 2: Block diagram of the proposed method

- Input Video: Input video is uploaded or live stream video is given as input.
- Video to Frames: This block converts the input video into image frames.
- Frames Histogram Difference: This block calculates the histogram differences between the present frame and the previous frame to find the abnormality in the activity.
- Threshold Comparison: This block compares the histogram differences of the frames and decides the activity.
- Output Activity Detection: This block gives the final output whether the activity is normal or abnormal.

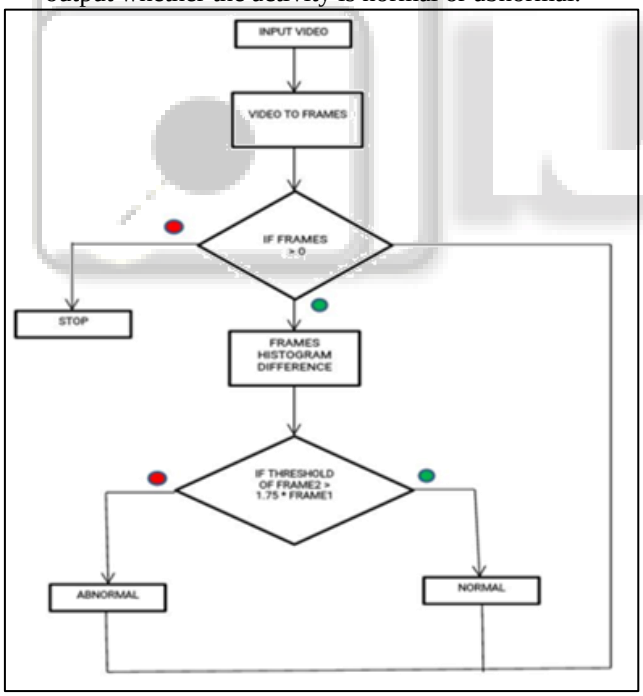

Fig. 3: Flowchart of the proposed system

The flowchart of the proposed system is depicted in the above figure 3. For each video frame Histogram differences is calculated by other frame. If the threshold of the frame 2 is greater than 1.75*Frame 1 then the activity is considered to be an normal activity or else it is been considered as an abnormal activity.

A. Histogram

A histogramis are presentation of approximate distribution of numerical data. It was developed by Karl Pearson. The first step to construct the histogram is to "bin" the range of values that divide the entire range of values into a series of intervals and count how many values will fall into each interval. The bins are normally stated as, non-overlapping intervals, consecutive of a variable. The bins must be adjacent, and are often of equal size.

A rectangle is erected over the bin with height proportional to the frequency the number of cases in each bin if they are in equal size. To display the "relative" frequencies a histogram may also be normalized. It then shows the proportion of cases with the sum of the heights equaling 1, that fall into each of several categories. Bins need not be of equal width; in such case, the erected rectangle is defined to have its area proportional to the frequency of cases in the bin. The vertical axis is then not the frequency but it's density of the number of cases per unit of the variable on the horizontal axis. Census bureau data is used to display the variable bin width as shown below in fig 4 and histogram difference is shown in fig 5.

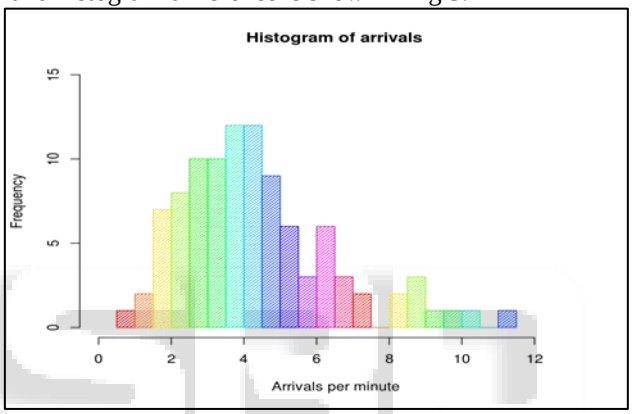

Fig. 4: Histogram frequency

The rectangles of a histogram touch each other to indicate that the original variable is continuous, As the adjacent bins leave no gaps.

Histograms gives a sense of the density of the underlying distribution of the data, and then often for density estimation: It estimates the probability density function of the underlying variable. The area which is used of a histogram for probability density is always normalized to 1.The histogram is identical to a relative frequency plot, If the length of the intervals on the x-axis are all 1.

A histogram is said to be a simplistic kernel density estimation, which uses a kernel to smooth frequencies over the bins. This represents a smoother probability density function, which will be in general more accurately reflecting distribution of the underlying variable. An alternative to plot the histogram is the density estimation, and is commonly drawn as a curve rather than a set of boxes.

Histograms are nonetheless preferred in applications, when their statistical properties need to be modelled. It is very difficult to describe mathematically about the correlated variation of a kernel density estimate, while for a histogram it is very simple where each bin varies independently.

An average shifted histogram is an alternative to kernel density estimation, which is very fast to calculate and gives a smooth curve estimate of the density without using kernels.

The histogram is one of the seven basic rules of the quality control.

A histogram is used for continues data, where the bins represent ranges of data, while a bar chart is a plot of categorical variables. Sometimes histograms are confused with bar charts, while other authors recommend that bar charts have gaps between the rectangles to clarify the distinction.

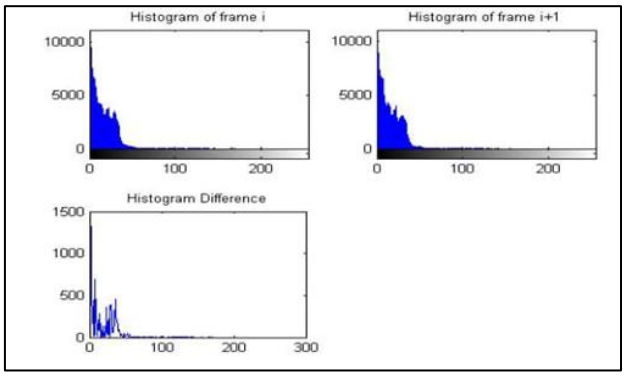

Fig. 5: Histogram difference

Threshold value for comparison between two frames is

Th2 > 1.75 * th1

Th1 = frame i - frame i+1

Th2= frame i - frame i+2

IV. RESULTS

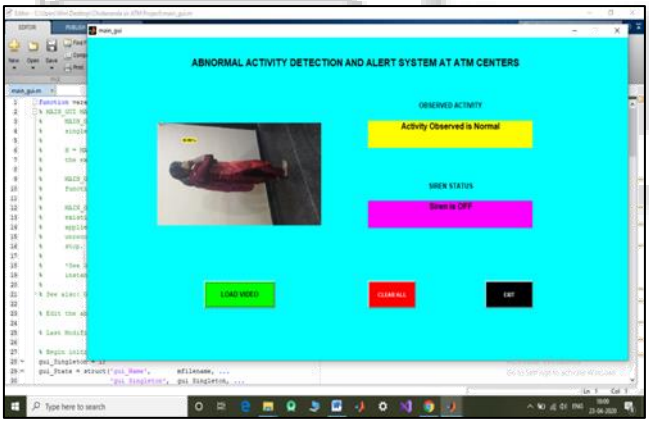

Fig. 6: Abnormal activity as input

Fig. 7: Abnormal activity detection and alert system

In the pre-processing stage of our system, we used the silhouette data that come along with the dataset. During the setup of the experiments, we configured different HMM models by changing the number of hidden states to different values among 3 to 8 and evaluated the recognition performance of the system separately for each model. Test results show that the best classification accuracy is obtained when the number of hidden states are configured using 5,6, and 7.Because the number of training samples for each action is low, the separation of the data into distinct train, test and validation sets is not reasonable. Therefore, we applied leave one-out technique in to evaluate the performance of the proposed solution. In this method, we set one video stream aside for testing and used the rest of the dataset for training. The obtained model is tested with the test stream. This process is repeated for all individual video streams in the dataset. The average accuracy is reported here. In order to compare the effectiveness of the modified MHI with silhouette based approach with HMMs; we performed additional experiments using the same features that are extracted using only the silhouettes and the modified MHIs. We also evaluated the method with a varying number of frames while forming the resample packages.

V. CONCLUSION

In this project, we have presented a system for security framework at ATM centres. In particular, this project recognizes the normal and abnormal activities at ATM Booth. The need of developing such security system increased due to increasing in crime rates at the ATM centres and also the lack of prevailing video surveillance system in the market. Overall accuracy of our system would have been higher if we had removed the transition frames between normal and abnormal activities. We have included this scenario since at real time this could not be eliminated. In research aspect the future scope of this paper is wide open. Our system is restricted to work for video only; our future aspect will be to focus on audio recognition also.

REFERENCES

[1] M. S. Scott, Robbery at Automated Teller Machines, US Department of Justice, Office of Community Oriented Policing Services, 2001.

[2] N. Sharma, “Analysis of different vulnerabilities in auto teller machine transactions,” Journal of Global Research in Computer Science, pp. 38–40, 2012.

[3] R. S. Shirbhate, N. D. Mishra, and R. P. Pande, “Video surveillance system using motion detection: a survey,” International Journal Advanced Networking and Applications, vol. 3, no. 5, pp. 19–22, 2012.

[4] IBN LIVE, 2014, http://ibnlive.in.com/news/bangalore-atmattack-womans-skull-fractured-still-in-icu/435190-62-129.html.

[5] NDTV, 2014, http://www.ndtv.com/article/cities/bangalore-this-brave-atm-guard-grabs-machete-from-robbers-hits-one-of-them-465052?curl=1422275786.

[6] MAN ATTACKED, January 2014, http://www.youtube.com/watch?v=RGaupYx2fpQ.

[7] R. Poppe, “A survey on vision-based human action recognition,” Image and Vision Computing, vol. 28, no. 6, pp. 976–990, 2010.

[8] U. Mahbub, H. Imtiaz, and M. A. R. Ahad, “Action recognition based on statistical analysis from clustered

flow vectors," Signal, Image and Video Processing, vol. 8, no. 2, pp. 243–253, 2014.

[9] W. Gong, J. Gonzàlez, and F. X. Roca, "Human action recognition based on estimated weak poses," EURASIP Journal on Advances in Signal Processing, vol. 2012, article 162, 2012.

[10] M. Paul, S. M. Haque, and S. Chakraborty, "Human detection in surveillance videos and its applications—a review," EURASIP Journal on Advances in Signal Processing, vol. 2013, article 176, 2013.

[11] W. Kim, J. Lee, M. Kim, D. Oh, and C. Kim, "Human action recognition using ordinal measure of accumulated motion," EURASIP Journal on Advances in Signal Processing, vol. 2010, Article ID 219190, 2010.

[12] M. Ahmad and S.-W. Lee, "Human action recognition using shape and CLG-motion flow from multi-view image sequences," Pattern Recognition, vol. 41, no. 7, pp. 2237–2252, 2008.

[13] J. W. Davis and A. F. Bobick, "The representation and recognition of human movement using temporal templates," in Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 928–934, June 1997.

[14] M. A. R. Ahad, J. K. Tan, H. Kim, and S. Ishikawa, "Motion history image: its variants and applications," Machine Vision and Applications, vol. 23, no. 2, pp. 255–281, 2012.

[15] M.-K. Hu, "Visual pattern recognition by moment invariants," IRE Transactions on Information Theory, vol. 8, no. 2, pp. 179–187, 1962.

[16] A F. Bobick and J. W. Davis, "The recognition of human movement using temporal templates," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 23, no. 3, pp. 257–267, 2001.

[17] Amato and V. D. Lecce, "Semantic classification of human behaviours in video surveillance systems," WSEAS Transactions on Computers, vol. 10, no. 10, pp. 343–352, 2011.

[18] Chen, R. Wu, Y. Ni, R. Huan, and Z. Wang, "Research on human abnormal behavior detection and recognition in intelligent video surveillance," Journal of Computational Information Systems, vol. 9, no. 1, pp. 289–296, 2013.

[19] P. Srestasathiern and A. Yilmaz, "Planar shape representation and matching under projective transformation," Computer Vision and Image Understanding, vol. 115, no. 11, pp. 1525–1535, 2011.

[20] S. Bourennane and C. Fossati, "Comparison of shape descriptors for hand posture recognition in video," Signal, Image and Video Processing, vol. 6, no. 1, pp. 147–157, 2012.

[21] G. Debard, P. Karsmakers, M. Deschodt et al., "Camera based fall detection using multiple features validated with real life video," in Intelligent Environments Workshops, vol. 10, pp. 441–450.

[22] E. B. Nievas, O. D. Suarez, G. B. Garcia, and R. Sukthankar, "Violence detection in video using computer vision techniques," in Computer Analysis of Images and Patterns, vol. 6855, pp. 332–339, Springer, Berlin, Germany, 2011.

[23] H. Wang, A. Finn, O. Erdinc, and A. Vincitore, "Spatial-temporal structural and dynamics features for Video Fire Detection," in Proceedings of the IEEE Workshop on Applications of Computer Vision (WACV '13), pp. 513–519, Tampa, Fla, USA, January 2013.

[24] S. Oh, A. Hoogs, A. Perera et al., "A large-scale benchmark dataset for event recognition in surveillance video," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR '11), pp. 3153–3160, Providence, RI, USA, June 2011.

[25] P. Rota, N. Conci, and N. Sebe, "Real time detection of social interactions in surveillance video," in Computer Vision—ECCV 2012. Workshops and Demonstrations, pp. 111–120, Springer, Berlin, Germany, 2012.

[26] D. Tosato, M. Farenzena, M. Spera, V. Murino, and M. Cristani, "Multi-class classification on Riemannian manifolds for video surveillance," in Computer Vision—ECCV 2010, vol. 6312 of Lecture Notes in Computer Science, pp. 378–391, Springer, Berlin, Germany, 2010.

[27] J. Arunnehru and M. K. Geetha, "Motion intensity code for action recognition in video using PCA and SVM," in Mining Intelligence and Knowledge Exploration, vol. 8284 of Lecture Notes in Computer Science, pp. 70–81, Springer International Publishing, Cham, Switzerland, 2013.

[28] B.-K. Bao, G. Liu, C. Xu, and S. Yan, "Inductive robust principal component analysis," IEEE Transactions on Image Processing, vol. 21, no. 8, pp. 3794–3800, 2012.

[29] R. D. Min, A. Angelo, and J. L. Dugelay, "Efficient scarf detection prior to face recognition," in Proceedings of the 18th European Signal Processing Conference (EUSIPCO '10), pp. 259–263, Aalborg, Denmark, 2010.

[30] G. Bradski and A. Kaehler, Learning OpenCV: Computer Vision with the OpenCV Library, O'Reilly Media, 2008.

[31] G. R. Bradski, "Motion detection using normal optical flow," U.S. Patent No. 6,654,483, U.S. Patent and Trademark Office, Washington, DC, USA, 2003.

[32] H. Witten, E. Frank, L. E. Trigg, M. A. Hall, G. Holmes, and S. J. Cunningham, WEKA: Practical Machine Learning Tools and Techniques with Java Implementations, 1999.

[33] C.-C. Chang and C.-J. Lin, "LIBSVM: a Library for support vector machines," ACM Transactions on Intelligent Systems and Technology, vol. 2, no. 3, article 27, 2011.

[34] A. Eisenberg and J. Melton, Java Programming Language.

[35] G. Bradski, "Programmers tool chest-the OpenCV Library-OpenCV is an open-source, computer vision library for extracting and processing meaningful data from images," Dr. Dobb's Journal: Software Tools for the Professional Programmer, vol. 25, no. 11, pp. 120–126, 2000.

