Intelligent Information System for Suspicious Human Activity Detection in Day and Night

J L Mazher Iqbal[1], S.Arun[2]
[1]Department of Electronics and Communication Engineering,
Vel Tech Dr.RR & Dr.SR Technical University, India
[2]Department of Electronics and Communication Engineering,
Madanapalle Institute of Technology & Science, Angallu, India

Article Info	ABSTRACT

The detection of human beings in a camera attracts more attention because of its wide range of applications such as abnormal event detection, person counting in a dense crowd, person identification, fall detection for care to elderly people, etc. Over the time, various techniques have evolved to enhance the visual information. This article presents a novel 3-D intelligent information system for identifying abnormal human activity using background subtraction, rectification, morphology, neural networks and depth estimation with a thermal camera and a pair of hand held Universal Serial Bus (USB) camera to visualize un-calibrated images. The proposed system detects strongest points using Speed-Up Robust Features (SURF). The Sum of Absolute Difference (SAD) algorithm match the strongest points detected by SURF. 3-D object model and image stitching from image sequences are carried out in the proposed work. A series of images captured from different cameras are stitched into a geometrically consistent mosaic either horizontally/vertically based on the image acquisition. 3-D image and depth estimation of un-calibrated stereo images are acquired using rectification and disparity. The background is separated from the scene using threshold approach. Features are extracted using morphological operators in order to get the skeleton. Junction points and end points of the skeleton image are obtained from the skeleton. Data set of abnormal human activity is created using supervised learning such as neural network with a thermal camera and a pair of webcam. The feature vector of an activity is compared with already created data set, if a match occurs the classifier detects abnormal human activity. Additionally the proposed algorithm performs depth estimation to measure real time distance of objects dynamically. The system use thermal camera, Intel computing stick, converter, video graphics array (VGA) to high-definition multimedia interface (HDMI) and webcams. The proposed novel intelligent information system gives 94% maximum accuracy and 89% minimum accuracy for different activities, thus it effectively detects suspicious activity during day and night.

Corresponding Author:

J L Mazher Iqbal ,
Department of Electronics and Communication Engineering,
Vel Tech Dr.RR & Dr.SR Technical University, Chennai, T.N, India

1. INTRODUCTION

Public security, safety and terrorism have been the chronic issue, a matter of serious concern for the public, government and entrepreneurs. A growing number of cases have been reported on child abuse, robbery and terrorism etc., threatening the general public. Though, there is a number of Clos Circuit Television (CCTV) cameras and dataset are available, nevertheless, monitoring and supervision of unwanted activities are becoming hectic and tedious. It requires more manpower to put untiring effort to keep everything under

surveillance. Therefore there is a chance of missing some events which are really suspicious because of human tendency. Detecting and recognizing suspicious activities are a complex, diverse and challenging area that have scope for modern day research [1]. Many countries are now employing video surveillance system. Detecting the abnormal behavior in real time will reduce major and minor crime rate. During the last decade, law enforcement agencies in Great Britain, France, Monaco, Spain and other countries have increasingly relied on Close Circuit Television (CCTV) surveillance to enhance public security. Siemens AG developed and installed the first video surveillance for observing launch of V-2 rockets in the year 1942 at Test Stand VII in Peenemunde, Germany. According to the survey presented in this report, an increasing number of U.S municipalities (mostly outside California) have begun using CCTV public video surveillance as a law enforcement tool to monitor public, schools and residential area. Many European countries are now employing public video surveillance as a primary tool to monitor population movements and to prevent terrorism. The United Kingdom (UK) in particular relies extensively on video surveillance as a tool to fight crime and prevent terrorism. CCTV image magazine had estimated the number of cameras in the United Kingdom where '1.85' million in 2011. Average one camera for every '32' people in UK, around '1.5' million CCTV cameras where installed in city centers, stations, airport and major retail areas in the UK. Typically CCTV surveillance systems where installed by British Telecom, using fiber-optic cables in strategically positioned down town business districts or in new shopping centers. The research organizations IMS (now part of IHM Inc. (NYSE:HIS) has released its fourth annual video surveillance trends for the year ahead white paper [2]. These prediction serves to provide insight into key trends and opportunities in the video surveillance industries. Video Image Retrieval and Analysis Tool (VIRAT) is in development by the US military's elite group of scientists, Pentagon's Defense Advanced Research Projects Agency (DARPA). VIRAT contract is announcement in May 2011 [3]. It describes specific human activities, military-intelligence video archives and systems. VIRAT system will be deployed into various military intelligence video archives and systems in Iraq. The tool is not for domestic surveillance programs. VIRAT focus on narrow targets, such as a doorway under video surveillance. The paper titled "Gait Recognition Using Radon Transform and Linear Discriminant Analysis" proposed by Nikolaos.V Boulgouris, IEEE Transactions on image processing (2007), presents the Gait Recognition Using Radon Transform and Linear Discriminant Analysis [4]. The paper titled "Real Time Human Activity Recognition System based on Radon Transform" proposed by Z.A. Khan and W. Sohn, IJCA special issue on Artificial Intelligence Techniques - Novel Approaches & Practical Applications" AIT (2011) presents human activity recognition system based on Radon Transform (RT), Principal Component Analysis (PCA) and Linear Discriminate Analysis (LDA), Artificial Neural Network [5]. The limitations of the LDA, PCA are small sample size problem (small no of very high dimensional samples), estimation unstable, high computation complexity, slow convergence, high variance is often introduced in the estimation for the zero. The paper [6] deals with fall angle comprising of fall recognition technique. The angle of fall can be defined as the slant among the base and the person from where it is certain that the human being will fall, or the incline among the base and the person's centroid. The angle of fall is in amid of '45' and '90' degree of angle, the human being is on foot, and while it is less than '45' angle of degrees, a fall can be concerned. The fall judgment as a result depends on the meaning of falling angle and on the manner and velocity of on foot. A limitation of using this method is of less accuracy, if the fall is towards the video camera it may not be recognized. The paper [7] shows depth estimation of inside Scanning Electron Microscope (SEM). To improve the accuracy as well as the rapidity of the method, it consider both autofocus and depth estimation as visual serving paradigms. Here depth estimation application shows in SEM only. Our earlier article [8], [9] discuss abnormality activity detection using Scale Invariant Feature Transform (SIFT) and Hidden Markov Model (HMM). The limitation of the paper [8], [9] is abnormal human activity detection is not accurate during night. The proposed system is the modification of the system proposed in our earlier article titled 'Abnormal Human Activity Recognition using Scale Invariant Feature Transform [8]. Thus public, government and private sector need accurate automatic surveillance monitoring system. The proposed intelligent information system use speed up robust features (SURF) to obtain the strongest point and sum of absolute difference (SAD) stereo images to match the features. The SURF and SAD is from our earlier article titled 'Real Time 3D Depth Estimation and Measurement of Un-calibrated Stereo and Thermal Images" [10]. The work done in the proposed area of research is analyzed with respect to document and year of publication, also with respect to document exist and country/territory using Scopus data base. The analysis are shown in Figure 1 and Figure 2. The analysis of document by country/territory (Figure 2) shows that out of '15' countries India stands in eleventh place in the proposed area of research. This shows that the scope at the National level is tremendous in the proposed area of research.
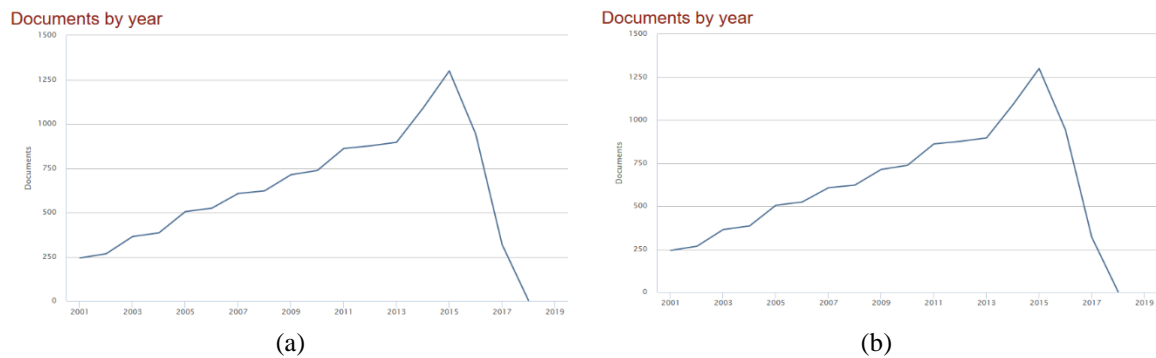
Figure 1. Analysis of Document and Year of Publication in the Proposed Area of Research
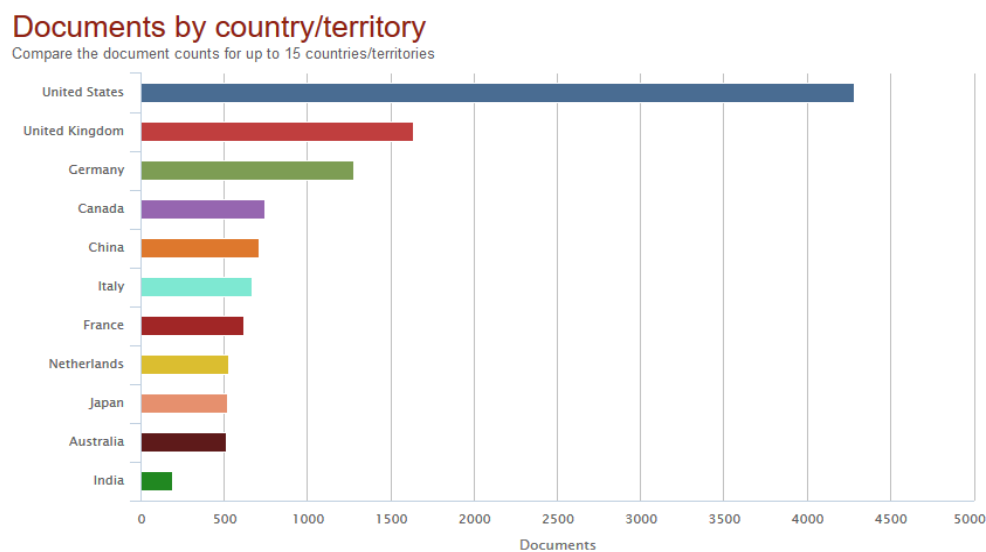(a) International Level (b) National Level
(Ref: www.scopus.com)

Figure 2. Analysis of Document by Country/Territory in the Proposed Area of Research
(Ref: www.scopus.com)

The gaps identified for the development of intelligent information system for abnormal activity recognition are less accuracy, inability of the system to detect action during night, estimation unstable, high computation complexity, slow convergence, high variance for suspicious human activity detection in day and night. The proposed work addresses this gap by providing solution to optimize the performance of the system using novel methods and algorithms such as SURF, SAD, image stitching, rectification, background subtraction, morphology, neural network and disparity. This paper presents an efficient Intelligent Information System that enables recognition of abnormal activity automatically during day and night using a thermal camera and a pair of webcam. Each abnormal activity in the video is classified into one of the predefined activity classes.

2. PROPOSED METHOD

A novel intelligent information system for identifying abnormal human activity recognition is shown in Figure 3, Figure 4 and Figure 5 show the data set creation using a thermal camera and various activities considered to be abnormal for the proposed dataset creation. Figure 6 shows the flow chart of the proposed novel intelligent information system. The video of different activities are captured using a thermal camera and a pair of webcam and is converted into frames. The color inputs are converted to binary for high-speed processing. Back ground image is created using background model. There are various challenges in developing a high-quality background subtraction algorithm used in the proposed system. Primary, it must be robust

against changes in lighting. Next, it should avoid recognizing non-stationary background objects such as moving leafs, rain, snow and shade cast by moving objects. The background image is created by setting a background and then deducting the current frame which contains the object of interest in motion. The incoming frame is subtracted from the background image using threshold technique in order to obtain foreground. A background subtraction algorithm attempts to identify moving regions by subtracting the current image pixel-by-pixel from a reference background. The pixels where the dissimilarity is more than a threshold, the background model classify as foreground object. A foreground object can be described as an object of attention, which helps in reducing amount of data to be processed. Foreground provide important information of the task under consideration. The proposed background subtraction algorithm efficiently obtain foreground for further processing.
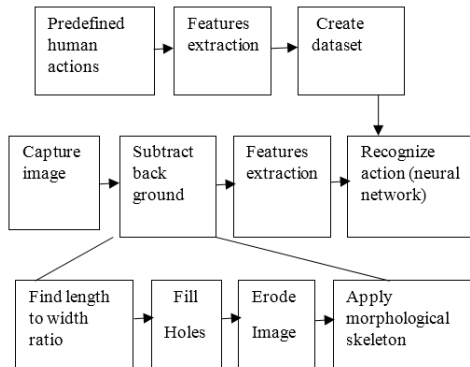

Figure 3. Intelligent Information System for Identifying Suspicious Activity

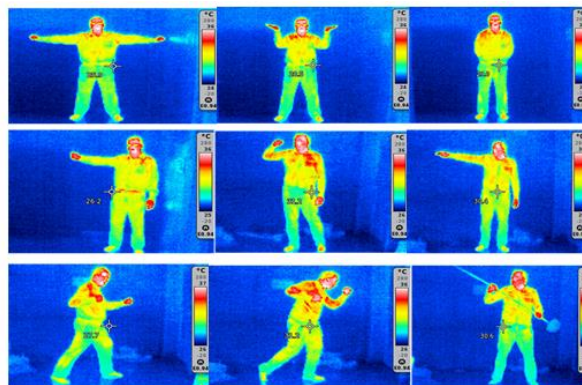Figure 4. Dataset Creation Using a Thermal Camera for Night Visio

Figure 5. Activities Using Thermal Camera

The foreground object is obtained using the equation

$$| \text{Frame } I_c - \text{Frame } I_b| > T_h \qquad (1)$$

where, frame I_c is the current frame, frame I_b is the background frame and T_h is the threshold value.

After creating a foreground pixel map, morphological post processing operations such as erosion, dilation and closing are applied in the proposed system to reduce the effects of noise and improve the quality of the scene. Before features extraction it is vital to find the length to width ratio of abnormal actions to detect suspicious human activity. Table.1 shows the length to width ratio of different suspicious human action to be detected during day and night. It is identified that the length to width ratio of human actions are 0.2 to 0.92. The length to width ratio will differentiate the human beings with other objects and it remove unwanted objects from the scene of interest. Subsequently various morphological preprocessing methods such as filling holes, erosion and morphological skeleton are applied to reduce the special effects such as clatter and it enhance the

quality of scene to be detected. The holes (region of black pixels bounded by white pixels) are filled using 'imfill' operator to remove the unwanted background spots in the region of action. Erode operation decreases the size of the image and remove small spurious bright spots. Corrosion in the proposed system shrinks objects by scratching away their borders. Features such as junction points and end point are extracted using morphological skeleton operator "bwmorph". Morphology skeletonization is the process of deleting unwanted pixels in order to obtain unit pixel width, it delete the pixels iteratively layer by layer. The junction points and end points are the feature vectors of the proposed system for supervised learning classifier. The green circles in the skeleton indicates end points and red circles indicates junction points. Neural network is used as classifier for classification of input patterns in the proposed system. Neural network is a system of interconnected "neurons" consist of elements which operates in parallel and can compute values from inputs through the network. In the proposed system, neural network is adjusted or trained, so that an input matches to specific target based on supervised learning. Data set of abnormal human activities are created using feature vectors obtained from a thermal camera and a pair of webcam for various abnormal activities shown in Figure 5. The classifier recognize a new test pattern based on feature vectors and created data set. If the input features matches with dataset the classifier recognize abnormal human activity. The system is tested for different abnormal activities such as, hands bending, hands up, standing etc. The behavior of the proposed system is evaluated using created dataset with high accuracy of around 90%.
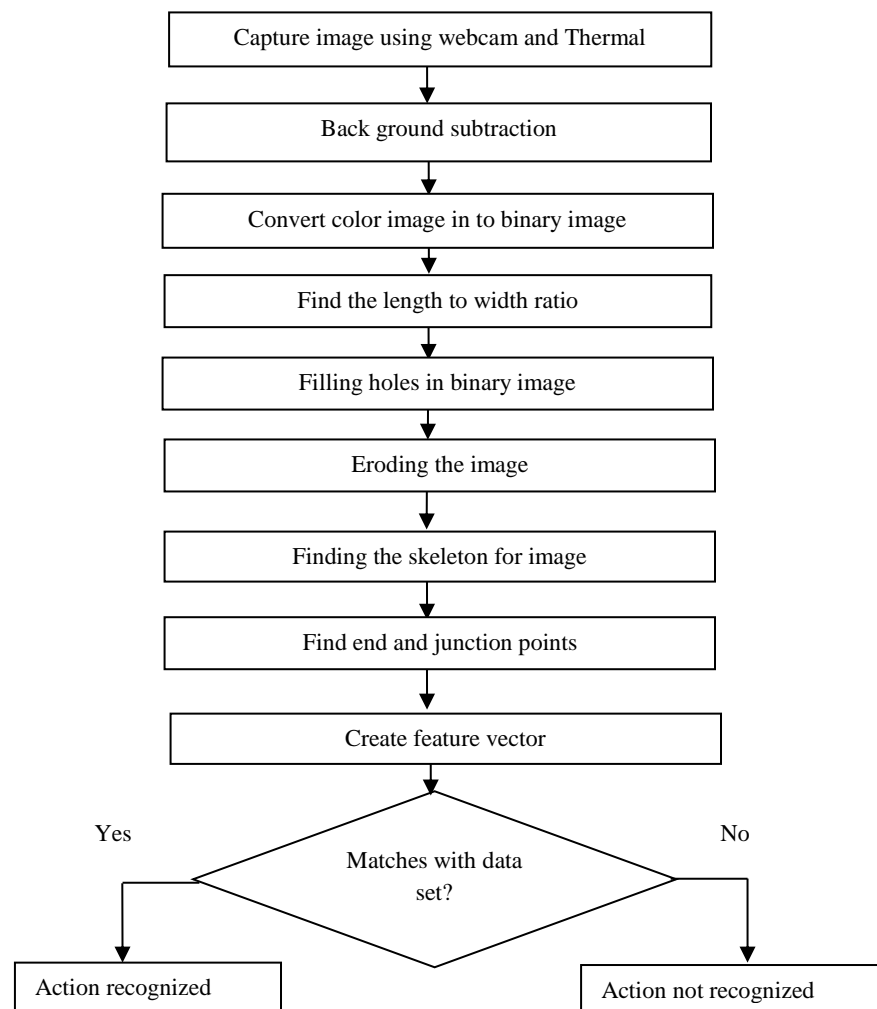
Figure 6. Flow Chart of the Proposed System Intelligent Information System
for Identifying Suspicious Activity

Table 1. Length to Width Ratio of Different Activities

No.of activities	Length	Width	Length/width
1	116	270	0.4296
2	63	228	0.2763
3	60	232	0.2586
4	134	242	0.5537
5	82	243	0.3374
6	73	241	0.3029
7	221	238	0.9285
8	218	238	0.9159
9	217	238	0.9117
10	86	237	0.3628

3.   RESULTS

The behavior of the proposed intelligent information system for identifying suspicious activity is evaluated using dataset created with different abnormal human actions. The experimental results of the proposed system are shown in Figure 7, Figure 8 and Figure 9. Figure 8 shows the accuracy of the proposed intelligent information system for identifying various activities. The proposed system gives 92%, 89% and 94% accuracy for activity '1', '2' and '3'. The proposed system gives 94% maximum accuracy and 89% minimum accuracy for various activities shown in Figure 5. Thus the proposed novel intelligent information system detects various suspicious activities during day and night effectively.
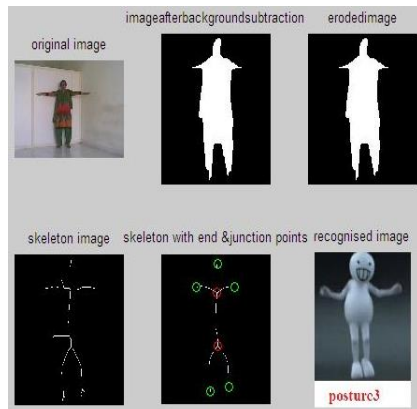

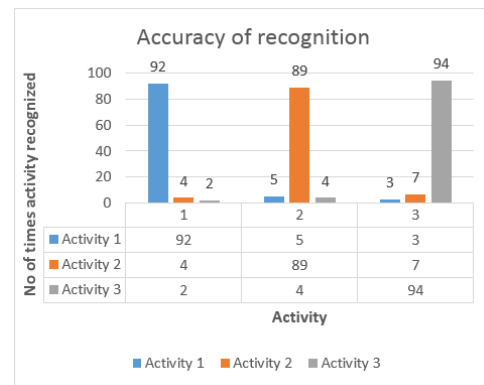Figure 7. Proposed Intelligent Information System

Figure 8. Accuracy of the Proposed Intelligent Information System for Identifying Different Activity
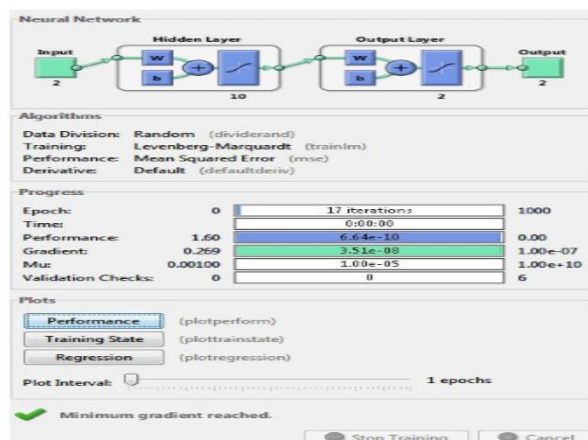

Figure 9. Neural Network Model for the Proposed Intelligent Information System

4.   CONCLUSION

In this paper, intelligent information system for identifying different human actions using background subtraction, rectification, morphology, neural networks and depth estimation with a thermal camera and a pair of webcam to visualize un-calibrated images is presented. The methods and algorithms are tested using various anomalous actions such as hands bending, hands up, standing etc. For object recognition, reliable background model set up is required to identify moving object and update the background. Various common moving object detection techniques are available such as optical flow, segmentation, temporal differencing and background subtraction method. The proposed project uses background subtraction method. The proposed work performs background subtraction for moving object recognition using threshold method. The algorithm is improved such that images with non-uniform background can also be used, this may enhance the human computer interaction. The moving object is tracked by calculating the area and centroid. Features such as junction points and end points are extracted using morphology skeleton. Supervised learning such as neural network is used to detect anomalous human action.

ACKNOWLEDGEMENT

REFERENCES

[1]   Lena Gorelick, Moshe Blank, EliShechtman, Michal Irani, Member, Ronen Basri, Member. "Actions as Space-Time Shapes" *IEEE Transactions on Pattern Analysis & Machine Intel*ligence, Vol. 29, No 12, pp.2247-2253,2007.
[2]   http://www.imsresearch.com/researcharea/Security_and_Fire/Video_Surveillance_and_VCA.
[3]   http://www.kitware.com/avss_2012.html.
[4]   Nikolaos V. Boulgouris, Member, IEEE, Zhiwei X. Chi. "Gait Recognition Using Radon Transform and Linear Discriminant Analysis"*IEEE Transactions on Image Processing*, vol. 16, no. 3, March 2007.
[5]   Arwa Darwish Alzughaibi "review on human motion detection based on background subtraction techniques, *IJCA.*, 2014.
[6]   V. Vishwakarma, C. Mandal, S. Sura. "*Automatic detection of human fall in video*", Lecture Notes in Computer Science, 4815:616-623, 2013.
[7]   Naresh Marturi, Brahim Tamadazte, Nadine Piat."Visual Servoing-Based Depth-Estimation Technique for Manipulation inside SEM" *IEEE transactions on instrumentation and measurement*, vol. 65, no. 8, August 2016.
[8]   J.L.Mazher Iqbal, J.Lavanya, S.Arun, "Abnormal Human Activity Recognition using Scale Invariant Feature Transform" *International Journal of Current Engineering and Technology*, Vol.5, No.6, PP 3748-3751, December 2015.
[9]   J.L. Mazher Iqbal, M. Suriya Parveen S. Arun. "Image stitching and 2D to 3D Image Reconstruction for Abnormal Activity Detection", *International Journal of Computer Applications*, Volume 133 – No.17, PP 1-7 January 2016.
[10] J.L. Mazher Iqbal, Shaik Shakir Basha "*Real Time 3D Depth Estimation and Measurement of Un-calibrated Stereo and Thermal Images*", International Conference on Nascent Technologies in the Engineering Field (ICNTE-2017), Jun 2017.