A KINECT-BASED REAL TIME VIDEO SURVEILLANCE FOR THE MOTION DETECTION AND TRACKING SYSTEM

K. Vijayalakshmi
Asst. Prof MCA Department Ethiraj College for Women
A. Susmitha
Research Scholar MCA Department Ethiraj College for Women

Abstract: Smart Video surveillance system stores a huge amount of sequential moving objects. This system assists examination of the video; focus to detect moving object and segment objects. Object classification step identifies detected objects into preened classes such as man, creature, vehicle, clutter etc. This paper focused to analyze the task of tracking and object detection in ATM at real time video surveillance using Kinect sensor with two major modules, namely, knife detection and fall detection based on k-means clustering algorithm. The Kinect captures both standard camera data and a three-dimensional depth map. At the receiver's end, the processed video, location identifier is send through MMS and the security; privacy issues in surveillance applications are applied.

Keywords: Kinect, K-means clustering, MMS, Cloud storage, GSM, DSA, LBS.

Introduction: In recent years, a massive number of cameras have been found out in the open places as a part of intellectual video surveillance framework. Such system monitors informal activities, protect people and property. In video surveillance system, technician or human worker is in charge of monitoring all tasks while watching 24 hour video originating from various cameras. Therefore, these systems are useless for huge thronged places as the number of cameras overreach the capacity of human expert. Video cameras are expensive in prior, but spontaneous retrieving, monitoring and verifying enormous amount of data is impractical, so from the domain computerized detecting and tracking people in banking system(ATM) is competitive in computer. This system overcomes the problem by using video analysis and moving object detection to be proceeded by tracking objects frame-to-frame results and take the variety of solutions into consideration: effective object detection, tracking, detect and identify fault in movement analysis with alarming situation and location based services(LBS) send as an Multimedia Messaging Service(MMS) to the end user [3] . The consecutive way to deal with moving object detection is done through background subtraction that has up-to date model detecting the present video from the background. There are various justification for tracking objects. The most essential application is people checking, object detection and security.

Human action recognition involves automatically detecting and analyzing human actions from the information obtained from sensors such as depth cameras, RGB cameras, range sensors, wearable inertial sensors, or other type sensors.

The Kinect originally designed to detect three dimension positions and to estimate human pose. In this paper, the video input object is created for the two streams: color video and depth video. The depth camera provides the detected person in gray level encoding (the depth information) and recognize through skeletal tracking of individuals which is a core component of the human detection. After the recognition of depth video and skeletal tracking, RGB camera stores the human as a color video in the server. Whenever the fraudulent image such as a person with knife or human fall inside the ATM is detected; the focal message alarm is send as MMS with the detected video to the police mobile with the use of GSM modem.

In this method Gaussian Mixture Model (GMM) is used for object detection and K-means clustering for Object tracking. By using Background Subtraction to detect foreground object in an image taken from stationary camera.

Distributed Server Architecture(DSA) with two servers is facile to setup in an Automated Teller machine for storing motion detection in cloud and retrieving the video for sending an alert from the cloud. Banking Alarm Systems are normally to keep bank secure from interruptions. GSM modem is a remote electronic gear acts as a dial-up modem. The dial-up modem receives the data through radio waves.GSM modem needs a SIM card from remote transporter for functionality.

Reserch Review:

MengChe Chuang, Jenq-Neng Hwang, Kresimir Williams, Richard Towler, 2014. : In this paper, they addressed the issue of tracking an object in a video its area in the frame and no other data. As of late, a class of following techniques called "tracking by identification" has been appeared to give encouraging outcomes at constant rates. These systems set up a discriminative classifier in an online approach to seclude the object from the background. This classifier bootstraps itself by utilizing the present tracker state to remove positive and negative cases from the present frame [5].

Chia-Hung Yeh, Chih-Yang Lin, Member, Kahlil Muchtar, Hsiang-Erh Lai, and Ming-Ting Sun, 2017.: This paper depends on the ideas of hysteresis Thresholding which alleviate cavity problems in foreground objects and movement remuneration to recover broken foreground objects such as slow motion and moving towards camera, which constitute spatial and temporal compensation, individually. Experimental results comes about demonstrate the current moving object identification techniques as far as accuracy, review, F-measure, and different estimations [1].

Rajalakshmi, A.Akash Kumar, S.Bala Kumar and N.Sailesh Khanna, 2017.: This paper implements a pixel wise background modeling which compares and detects the moving object and segment objects supported motion information and alert the end user. In order to detect efficient moving object a contiguous outliers in the generic algorithm which is used for efficient object detection [6].

Adkinson-Orellana, D. A. Rodríguez-Silva, F. J. Gonzalez-Castano, 2012.: This paper gathers multimedia streams created by observation cameras, streamlines their transmissions as indicated by network condition and stores them in a distributed storage framework in an effective and secure way. The dynamic sending of particular cloud processing servers require programmed switching between cloud suppliers. The framework depends on client and handling server buffers to limit the loss of information, permits to manage with traffic peaks and capacity benefit irregularity [7].

Consuelo Granta, Ragou Ady, and Joseph Salini, 2013.: In this paper, combining a multi model human detector based on sensors in 3D space is a fundamental undertaking for assisted by comparing with dynamic posturography platform applications. The Pearson's correlation coefficient is computed to consolidate three dimensions (RGB) histogram. It evaluates the benefit of joining color and depth information of multiple cameras recorded with Kinects which results the significant improvements in challenging scenes [8].

M.Dennish Ammu, P.Puviyarasi , S.Sathish Kumar, 2017.: The android alert system reduces the crime or illegal activities in the monitoring site. The object detection using the background subtraction techniques reduces the time effort greatly and it will not affect by any sensitive light changes. The capability of mobile monitoring extends to detecting foreground objects under all reasonable lighting conditions, tracking moving objects as they move along their trajectories, measuring the performance of the application as opposed to the ground truth of the video, handling occlusions and noises, projecting the object separated from its background [9].

Mandakini A. Mahale and H. H. Kulkarni, 2017.: The abandoned object detection Bank representation provides different methods to detect and track objects. Kalman Filter is used for tracking objects in an image taken from a camera and computes the following: detection count statistics (per object category) for each grid cell in a spatial pyramid (entire image, 2→2, and 4→4 grid): the sum of scores of detections within that cell (above a certain threshold), the number of detections, and a single bit that indicates whether or not there is a detection within in that cell [10].

Christoph Busch and R. Raghavendra, 2014.: This paper implements novel Presentation Attack Detection (PAD) algorithm based on both local and global features. The local features are extracted to capture the artefacts in eye and nose region contributed due to the presence of the 3D mask while global features are used to extract the micro texture components using Binarized Statistical Image Features (BSIF). Finally the comparison scores obtained using local and global features on both 2D color and depth image using weighted sum rule are fused [11].

Table I- Analysis of Various Techniques for Smart Camera Security System

S.No	Author	Year	Methods	Pros	Cons
1	MengChe Chuang	2014	Piecewise parametric motion segmentation	Doesnot require pre-defined pattern templates.	The model is based on the assumption of small motion. This means that motion should not exceed four or five pixels between frames. Struggles to identify non-moving humans.
2	Chia-Hung Yeh	2017	Frame Differencing	Simplest background subtraction	Difficult to implement in realtime system
			Gaussian Mixture	Low memory management	Cannot copeup objects with noise.
3	D. A. Rodriguez-Silva	2012	Distributed Storage	collects images generated by surveillance cameras, optimizes their transmissions according to network condition and stores them in a cloud storage system	Based on today's market a cloud computing solution is not less expensive compared to purchasing the hardware and supplying power.
4	M. Dennish Ammu	2017	Connected Component Analysis	assign a unique label to all pixels of each connected component (i.e., each object) in a binary image.	The major problem caused by the merging step is that all the previously scanned pixels need to be relabeled with one unique label per object.
5	R. Raghavendra	2014	Presentation Attack Detection(PAD)	detect and mitigate the 3D mask attacks both local and global features from the captured face image.	False positive indications wrongly categorize the impairing the efficiency of the system and not preventing from the security breach
ANALYSIS OF DEVELOPING SYSTEM					
6	Background Subtraction			This method is simple and easy to realize, and accurately extracts the characteristics of target data.	It is sensitive to the change of external environment, so it is applicable to the condition that the background is known.
7	Temporal differencing			This is very adaptive to dynamic scene changes.	Sensitive to dynamic changes. Need background frame with still objects.
8	Mask Motion Object Detection			Adopted to reduce noise and smooth object edge and get good detection result by optimizing memory.	Reduce computational complexity. Experiments with specific problem caliberate the parameters of the algorithm.

From the above review in real time video surveillance alert system with kinect object detection and frame segmentation, the moving object fragment detection is tested using primary k video frames. K-means cluster algorithm are handled by the background subtraction module to track the target object by checking the differences between the observed image and the predefined background model. This system uses depth map, skeletal tracking and color image obtained by kinect sensor to identify knife and helmet detection. It uses background subtraction and k-means clustering algorithm for comparison of images. After the comparison operation the threshold value is specified to check the object detected exceeds the limit. The sectioned areas are distinguished by consolidating two unique techniques: K-means tracker is used to solve the background interfusion by introducing the negative samples as well as positive samples into object tracking. The discern object recognition is stored in switching between cloud processing servers.

Existing Work:

Video Capturing: A static camera observing a scene is a common case in a video surveillance application. Detecting moving objects is an initial step toward high-level image processing. The most typical method for this problem is background subtraction.GMM is used to model the background.

Gaussian Mixture Model: The Gaussian mixture model was first proposed by Grimson and Stauffer[13]. Each pixel in the scene is modeled by a *K*-Gaussians mixture model (GMM).Taking into consideration lighting changes, scene changes and moving objects, the model using Gaussian mixture is more reliable and adaptive. The probability of observing the current pixel value is:

$$P(x_t)=\sum_{j=1}^{k} \omega_{j,t} * \eta(x_t, \mu_{j,t}, \Sigma_{j,t})$$

where x_t is the current pixel value at time t, *K* is the number of Gaussian distributions, $\omega_{j,t}$ is the weight estimation of the j^{th} Gaussian in the mixture at time t, $\mu_{j,t}$ and $\Sigma_{j,t}$ are the mean value and covariance matrix, respectively, of the j^{th} Gaussian in the mixture at time t, and η is a Gaussian probability density function (*pdf*).

After the model is built, each incoming pixel of the following frames is compared with the existing model components. In the case that the input pixel fits one of the weighted Gaussian distributions, which means that its pixel value is within standard deviations, the update process will be invoked to fine-tune the corresponding model; otherwise, the distribution, which has the lowest weight is replaced with a new distribution using the current incoming pixel as its mean value, an initial high variance, and a low prior weight.

To select the best Gaussians for each pixel, the *K* distributions are sorted based upon the value ω_k/Σ_k. Only the first *B* distributions are selected as the background model of a pixel for the scene and denoted as follows:

$B = \arg\min_b \{\sum_{k=1}^{b} \omega_k \mid \sum_{k=1}^{b} \omega_k > T', 1 \leq b \leq K\}$,

where *T'* is a predefined threshold and usually set to approximately 90%, ω_k is the weight parameter of the k^{th} model component and *b* indicates the number of background distributions.

In the updating process, the weights of *K* Gaussian distributions will be changed as follows:

$\omega_{k,t} = (1-\alpha)\, \omega_{k,t-1} + \alpha(M_{k,t})$,

where α is the learning rate and Mk,t is 1 for the matched distribution and 0 for the unmatched distributions. In addition, weights of distributions should be renormalized. If the new pixel matches a Gaussian distribution, the values of mean and variance of this distribution are updated as follows:

$\mu_t=(1-\rho)\, \mu_{t-1} + \rho x_t$

$\rho = \alpha\, \eta\, (x_t \mid \mu_t, \Sigma_t)$

Tracking Algorithm: Once the object has been recognized then it can be followed along its path.

Kalman Filter: A Kalman filter is utilized to appraise the consecutive framework where the state is assumed to the Gaussian distribution. The Kalman filter is a recursive prescient filter that depends on the utilization of space methods and recursive algorithms. Object tracking is performed by predicting the object's position from the past data and checking the presence of the object at the predicted position. The Kalman filter is a set of mathematical equations that gives a proficient computational (recursive) intends to gauge the condition of a procedure in several perspectives: it supports estimations of past, present, and even future states, and it can do the same even when the precise nature of the modelled system is unknown. The Kalman filter estimates a process by utilizing a type of feedback control. The equations for Kalman filters fall in two groups: time update and estimate update equations. Kalman filtering is composed of two step: 1) Prediction 2) Correction

State Prediction $Xpred_k = A^* X_{k-1} + B^* U_k + W_{k-1}$ (1)

Error Covariance Prediction $Ppred_k = A^* P_{k-1}{}^*A^T + Q$ (2)

k-1 is a vector process state.

$Xpred_k$ is vector representing predicted process state at time k.

$Ppred_k$ is predicted error covariance at time k.

Drawback: When a moving object in a real time video surveillance suddenly stopped for a while, it is absorbed into the background because it is not a moving object anymore. This causes the foreground object to be broken up and this issue is hard to solve using GMM and gives a false result while implementing in real time.

The existing system has been applied for colour and gray scale video imagery but moving object when comes near or far from video surveillance system than it is disappeared when it convert to a specific resolution.

Kalman filter is not suitable for real-time processing because of long computational time and able to represent only in Gaussian distribution.

The performance of the system cannot implement in real time because due to the storage capacity is high the system doesn't keep track of the previous footage and always need a human operator to save the image sequence in another device when the storage exceeds the file size and check whether the object is captured without cavity problem. It also sends an alert to the end user whenever a human has been captured in the video and it stores only unique images; so, that when the fraudulent appear in a sudden; the server doesn't have a proof of a continuous detected video.

Developing System:

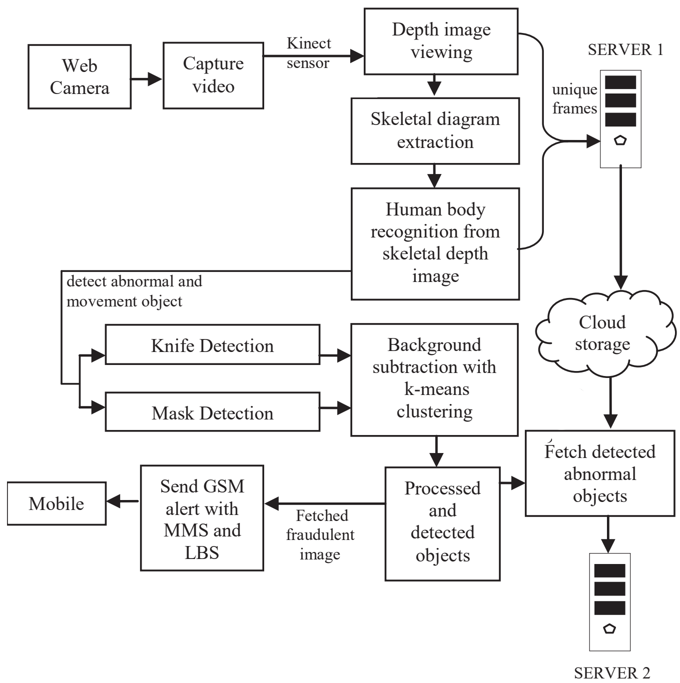

Fig. 1: Architecture of Smart Camera Alert System

The above diagram shows the fraudulent image detection with effective secure storage and detected object retrieval using cloud.

Apprehend Video: After the video capturing is done; Depth data and RGB data from the two Kinect cameras were recorded for the identification of fraudulent sequences. This dataset also consists of moving data related to the normal activities of walking, sitting down and lying in order to evaluate the performance of the motion object detection. After the recognition of depth mapping and skeletal tracking the human body movement captured by the camera inside the surveillance area will be send to the server1 and store the frame sequence in cloud.

Kinect Sensor: The Microsoft kinect was launched in the year 2010.The kinect sensor (Fig.2) consists of two cameras: 3D Depth camera, infrared sensor, RGB camera and four microphone arrays. The depth information can be easily adopted to detect moving objects. The depth sensor comprises the IR projector and the IR camera. The IR projector projects a known light speckle which e is invisible to the color camera but can be viewed by the IR camera. RGB camera delivers three basic color components of the video.

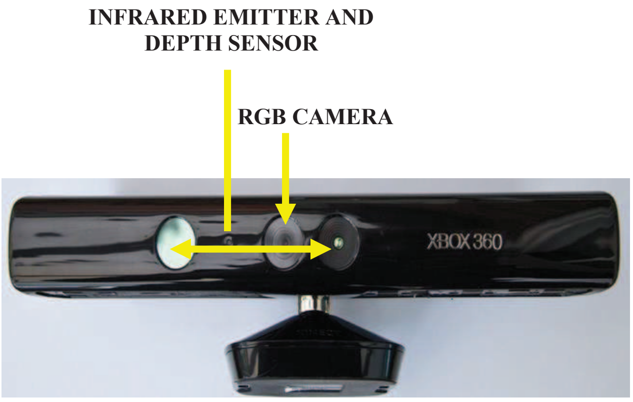

Fig. 2: Kinect Sensor

Depth Image Viewing and RGB Camera:

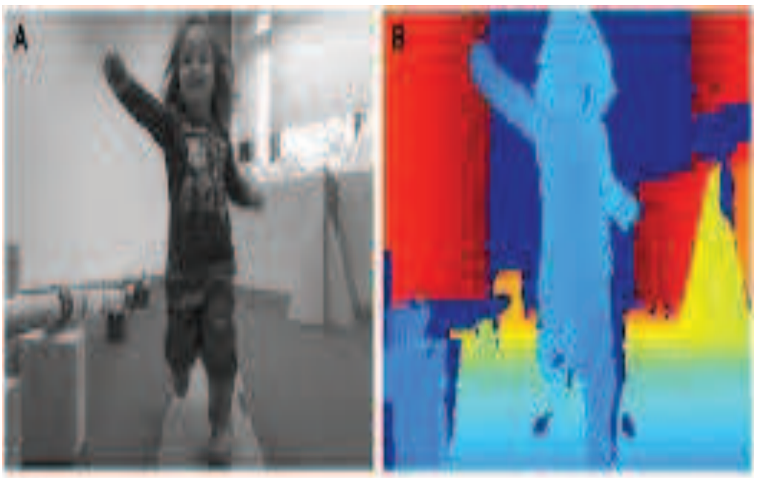

Fig. 3: Kinect sample output

In Fig. 3 there are a few samples showing how Kinect output looks like. From left to the right: depth camera and infrared sensors.
Create the VIDEO INPUT objects for the two streams

Color Video = video input('kinect',1);
Depth Video = video input('kinect',2);

The Frames Per Trigger property specifies the number of frames the video input object acquires each time it executes a trigger using the selected video source.
Color Video. Frames Per Trigger = 100;
depth Video. Frames Per Trigger = 100;
When the value of The Trigger Repeat property is set to inf, the object keeps acquiring frames until an error occurs or you issue a stop command.
Depth Video. Trigger Repeat = inf;

Skeletal Diagram Extraction: When the person is in front of the Kinect, his/her skeleton is detected and the 3D positions of all his joints are provided. The below diagram shows the centre axis lines for human body, but when region detection is needed the central point doesn't establish for a clear result .

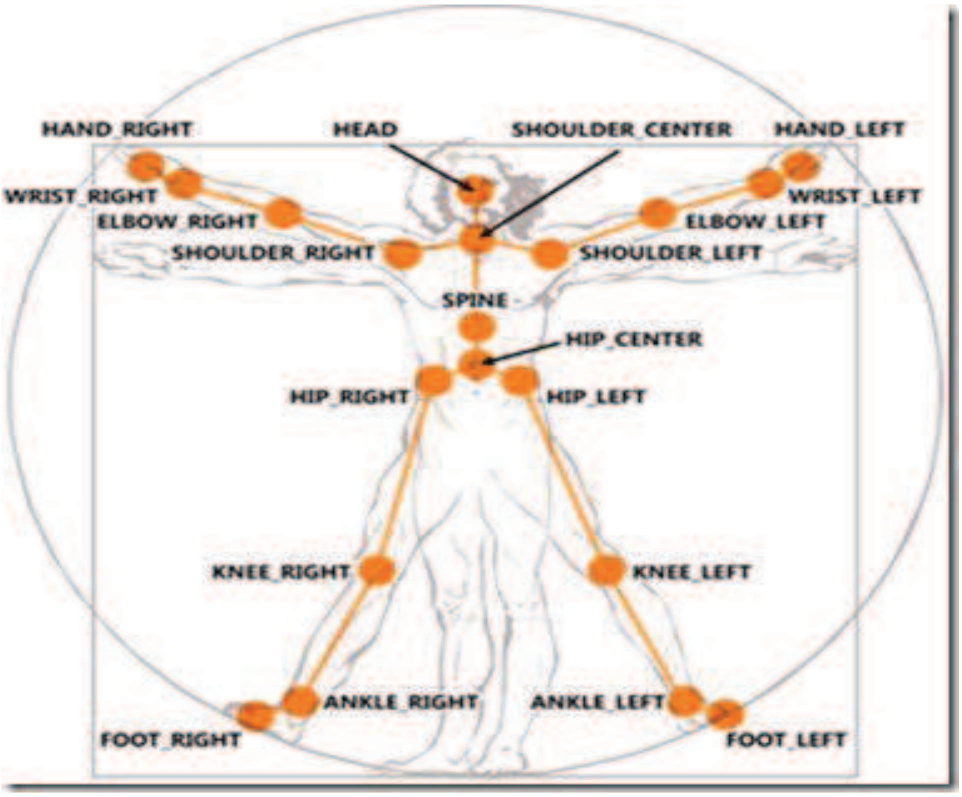

Fig. 4: Kinect Skeleton Tracking Points

Retrieve the frames and check if any Skeletons are tracked
[frame Data Color] = get data(color Vid);
[frame Depth, time Depth, meta Depth] = get data(depth Video);
View Skeletal Data from depth Meta Data and track skeletons
if sum(depth Meta Data. Is Skeleton Tracked) > 0
tracked Skeleton = find(depth Meta Data. Is Skeleton Tracked)
Find number of Skeletons tracked
N Skeleton = length(tracked Skeletons);
Plot the skeleton
util_skeleton Viewer(joint Indices, depth Map, n Skeleton);

Background Subtraction: In this method, the best way to calculate distance of moving objects is by determining the minimum and maximum distance, the background can be removed. For Min depth and Max depth, fix value were used as following statement. Min depth = Fix value
Max depth = Fix value

Codebook Background Subtraction Model is based on a quantization/clustering method [14]. The background model for each pixel is composed of a codebook consisting of one or more code words. Codebook algorithm consists of three different stages: construction of the initial codebook, foreground detection and model maintenance. Model Construction Given a set of N frames, a training sequence, S, is used for each pixel consisting of N RGB vectors:
$S = \{v_1, v_2, \ldots v_N\}$Initially, each pixel has an associated codebook, represented as $C = \{c_1, c_2, c_3, \ldots c_L\}$, consisting of L code words. The number of code words for each pixel may be different. Each code word, c_i , i = 1...L;, consists of an RGB vector,
$vi = (\overline{Ri}, \bar{G}i, \bar{B}i)$, and a six-tuple $aux_i = (I^i_{min}, I^i_{max}, f_i, \lambda_i, p_i, q_i)$. The tuple, aux_i , contains intensity values and temporal variables as described below:

- vi = ((Ri)¯ ,G¯i ,B¯i) , average value of each color component.
- I_min^i, I_max^i, minimum and maximum brightness, respectively, of all pixels assigned to codeword, ci .
- fi , the frequency with which codeword ci has been accessed.
- λi , the maximum negative run-length (MNRL), defined as the longest interval of time during which codeword ci has not been updated.
- p, q, the first and last updating access times of codeword ci . Some of the values of the codeword (λi , p, q) are only used to deal with the presence of foreground objects during the construction. In this paper the focus is on the use of color and brightness variables.

Algorithm 1: Algorithm for codebook construction
$C \leftarrow \phi$
for $t = 1 \rightarrow N$ do
$xt = (R, G, B), I \leftarrow \sqrt{R^2 + G^2 + B^2}$
Find the codeword, cm, in C matching to x_t based on two conditions:
    (a) colordist(x_t , v_m)
    (b) brightness (I^i_{min} , I^i_{max}) = true
    if $C = \phi$ or there is no match then
    {Create new codeword and add it to C}
    Else
    {Update matched codeword}
    end if
end for

K-means Clustering Operation: After the comparison operation of background modeling; k-means clustering algorithm classifies a pixels from each frame and undergo pattern recognition process.
Algorithm 2: (K-means clustering)
begin initialize n, c, $\mu_1, \mu_2, \ldots, \mu_c$
do classify n samples according to nearest
recompute μ_i
until no change in μ_i
return μ_1 , $\mu_2, \ldots, \mu_c$
end
n denotes the desired number of patterns
c implicit the desired number of clusters
μ_i is the average of the sample from the i^{th} class

The basic aim of this approach is to segment human automatically using the K-means clustering technique. The introduced framework of moving object segmentation operates in five steps as follows:

Step 1: Read the input image of detected object.(i.e) the human with knife or human felt down due to some issue.
Step 2: Classify abnormal objects using K-means Clustering. To measure the difference between two images(normal and abnormal images), Euclidean distance metric is used.
Step 3: Label Each Pixel in the Image from the Results of K-means. For every pixel in our input, K-mean computes an index corresponding to a cluster. Every pixel of the image will be labeled with its cluster index.
Step 4: Finally, generate RGB Images that Segment the objects which has the abnormal identification.

Abnormal Object Detection: After the merging process the abnormal objects can be detected by the following techniques: The background subtraction and k-mean algorithm is implemented in knife and fall detection.

Knife Detection: It is difficult to recognize the offender and the defender automatically during processing because of the dynamics of such events. Therefore, the two regions are watched for every human outline found in the picture (every human silhouette is thought to be both a potential.

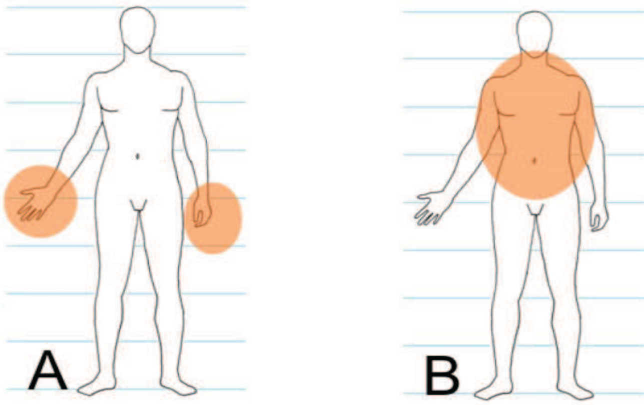

Fig. 5: Areas Where A Knife May Appear

The next step was to convert the picture into its numerical portrayal [12] by utilizing a sliding window mechanism to discover parts of images that contain highlights that are characteristic for knives. This way, the system can able to examine the approximate position of the knife in an image and no need to detect the knife's edges, which isn't inconsequential when pictures with a variable and non-homogenous background are considered.

The homogenous texture are computed using mean energy and energy deviation in each of 30 frequency channels.
The database consists of two classes of images:
From the Fig.5 A is taken as positive example and B is taken as negative example.

- Positive examples (PE): A knife held in a hand is noticeable in the picture. Only a knife held in a hand is considered to be a perilous circumstance. We consider a knife not being held by a man to be less dangerous. It can likewise be effortlessly discarded amid preparing or result in numerous false cautions.
- Negative examples (NE): A knife does not show up in the picture. NE dwarf PE to cover as many cases as possible. NE images were taken under comparative conditions as PE images.

Fall Detection: The skeleton outline extraction of the human body before the Kinect sensor influences utilization of both depth image and also the ordinary camera of the Kinect sensor for its exact three Dimensional vertical state. Change in its state is likewise quickly perceived by the IR sensor pair. The step by step methodology incorporates skeleton extraction, merging of depth image and skeleton, human body recognition, converting over and comparing pixels of RGB and Depth image, Monitoring movement condition of body and fall detection. Our initial step removes skeleton chart of the human body which is before the Kinect sensor. In the second step, the camera outline is made for the picture and profundity picture will be produced. It is the size of profundity of each protest in the directions concerning vertical separation from the Kinect. In the third step, the separated skeletal picture and depth images are converged into single object image. The global coordinates are then incorporated into the resultant picture. In this way, the depth image regarding global three dimensional directions is made for each key edges. In the fourth step, the human body is perceived and given constant depth portrayal for it, irrespective to other objects in the frame. This is precisely coordinated with the skeleton chart for consistency in recognition. After this progression, its time has to convert the RGB size of pixels into standard color imaging pixels of the depth image. This gives the perfect motion state of calculation to identify fall in the robust strategy.

Simulation Results:

Fig.6: Depth Stream Captured Image

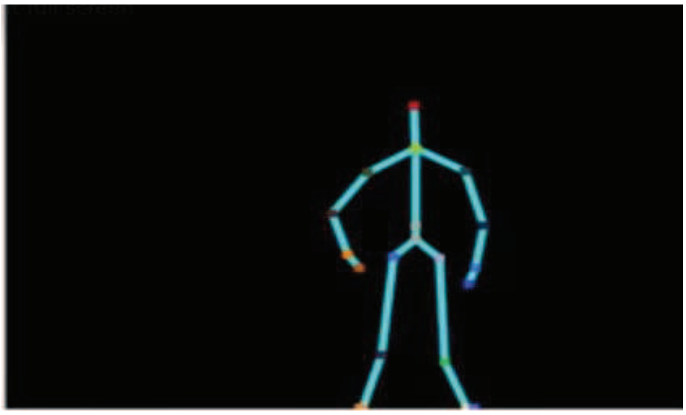

Fig.7: Skeletal Tracking (Rendered If Full Body Fits The Frame)

Fig.8: RGB Image When An Human Is Normal

Fig. 9: RGB Image When An Human Commence To Fall On The Floor

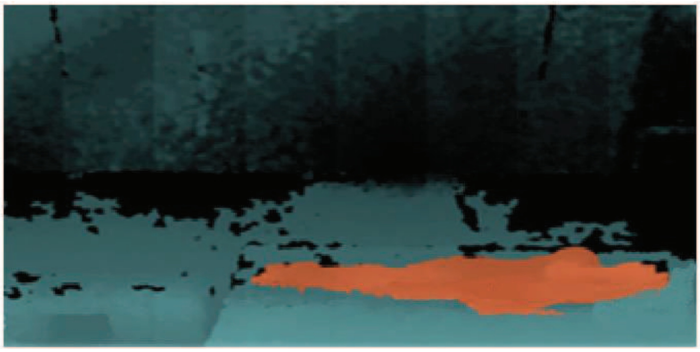

Fig. 10: The Human In An Abnormal Condition Alert The End User

Fig. 11: Alert Admin by Sending the Tracked Location

Alarm fraudulent image: When the detected object is processed; the detected data stored in cloud is fetched from the server 1 to the server 2.So that the server 2 sends the fraudulent image as an alarm to the admin.

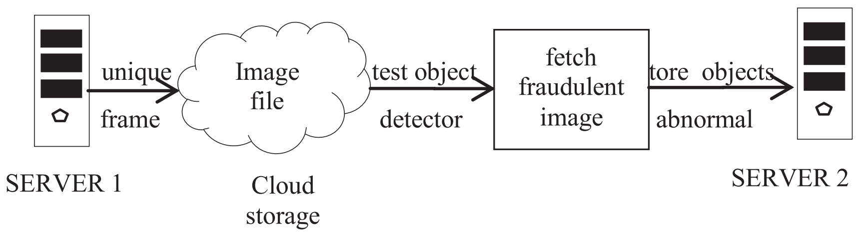

Fig.12: Flow of Suspicious Object Detection

GSM Modem: A GSM modem is an especially planned sort of modem that gets a subscriber identity module (SIM) card and works over a commitment to a portable administrator simply like a mobile.

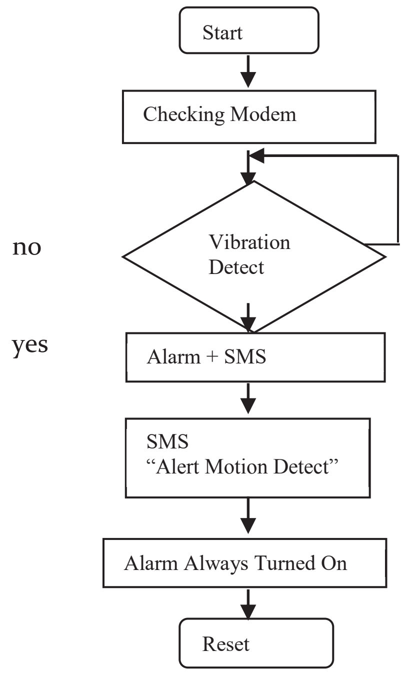

Fig.13 Flowchart of Surveillance System using GSM

The operational of GSM modem depends on commands, the commands dependably start with AT (which implies ATtention) and closures with a character.

It demonstrates that the GSM modem is controlled by the microcontroller that sends signals to the GSM to receive and transmit messages. In this framework, if the suspicious movement happened in bank or fraudulent attack has occurred, a caution indication send through SMS message to the admin's mobile; immediately and automatically by means for GSM modem. When receiving a predefined message with location based service (LBS), the owner can know where the attack has occurred with the detected video frame send to his/her phone and take an immediate action.

Conclusion and future Enhancement: This paper presents a widespread review of visual surveillance systems describing its phases of object detection, and tracking, semantic decision. Object detection techniques like background subtraction and k-means clustering for knife and fall detection are briefly described. Background subtraction technique is easy to implement with less calculation and generally used for real time applications.

Then, a unique approach brings into attention an uncommon classification that considers that humans are different than normal objects, therefore we try to use different methods for segmentation. These techniques helps in easy access of the images. However characteristics of each technique with their limitations are also described in this paper. Finally, alert the user sending multimedia SMS and Location tracking by using GSM is processed.

In implementation phase, a complete test with the improvements will be made further in the system and intend to implement in a widely accepted system in almost all banking sectors with all the required object detection techniques. This paper also plan to integrate both algorithms into a single solution, while further focusing on reducing false alarms and automatic determination of detected objects more accurately.

References:

1. Chia - Hung Yeh, Chih - Yang Lin, "Three – Pronged Compensation and Hystersis Thresholding for Moving Object Detection in Real - Time Surveillance ," IEEE Transaction on Electronics, vol.15, no.2, 2017.
2. Neetha Mary Paul, Ariya T. K , "Object Detection in Video Surveillance ," IJETT, 2014.
3. Michael Zablocki , Dariusz Feejlichwski, "Intelligent Video Surveillance System for Public Places, "vol.8, no.4, pp. 13-27, 2014.
4. Bian Junxia, Yin Jianqin, Wei Jun, Zhang Ling, "Hand Detection Based on Depth and Color Information of the Kinect," IEEE, 2015.
5. MengChe Chuang, Richard Towler, "A Variation Approach to Piecewise Parametric Motion Segmentation," IEEE, 2014**.**
6. S.Rajalakshmi, N.Sailesh Khanna, "Java Based Video Surveillance System Using Frame Separation In Real-Time," IJETT, 2017.
7. D. A. Rodríguez - Silva, Armino Franco, "Video surveillance based on cloud storage," IEEE, 2012.
8. Consuelo Granta, Ragou Ady, and Joseph Salini,"Aframework for robust Person detection asnd tracking ," 2013.
9. M.Dennish Ammu, P.Puviyarasi ,Sathish Kumar,"Surveillance of Object Motion Detection and Alert Using Android ,"IJIRSET, 2017.
10. Mandakini A. Mahale and H. H. Kulkarni, "Abandoned Object detection in Surveillance Video, " IJESC, 2017, 7, 11.
11. R. Raghavendra and Christoph Busch, "Novel Presentation Attack Detection Algorithm for Face Recognition System," 2014.
12. Michał Grega, Andrzej Matiolanski, Piotr Guzik and Mikołaj Leszczuk, "Automated Detection of Firearms and Knives in a CCTV Image," Sensors 2016, 16, 47.
13. C. Stauffer and W. E. L. Grimson, "Learning patterns of activity using Real - time tracking," IEEE Trans. Pattern Anal. Mach. Intell. , vol. 22, pp. 747-757, Aug. 2000.
14. Kyungnam Kim , Thanarat H. Chalidabhongse, "Background Modelling using Codebook Construction," 2004.

***