International Journal of Engineering Research and Development
e-ISSN: 2278-067X, p-ISSN: 2278-800X, www.ijerd.com
Volume 12, Issue 11 (November 2016), PP.10-16

Survey on Artificial Intelligence Techniques for Images Captured in A Body Worn Camera.

Sanjana C Mallik[1], Sharanitha Harish[2]
[1]BMS College Of Engineering, VTU, India.[2]BMS College Of Engineering, VTU, India.

Abstract:- Body worn cameras have carved the niche for efficiency in the identification of the miscreants. This paper surveys the various optimal methods for texture analysis and image classification from the video frames acquired from the body worn camera. The recommended configuration is a low powered, always-on camera, substantiated under the section II of related works. The frames would then undergo background subtraction which consists of background modelling and foreground detection which ultimately attenuates noise. This process proceeds for human detection followed by posture detection. The feature extraction is obligatory to zoom in on the posture of the silhouette. The multiple options for the same are surveyed such as gray level co-occurrence matrix(GLCM),texture feature, histogram feature, gabor feature, wavelet feature and region feature methods. The terminating phase consists of pattern recognition steps up the efficiency in image classification as it is equipped with copious algorithms and problem solving approaches. The prime division consists of supervised and unsupervised learning using artificial intelligence. Supervised learning consists of the following: Artificial Neural Network(ANN),Decision trees, Minimum distance, Bayesian network and Support Vector Machines(SVM).Of the unsupervised methods, the following approaches are explored: Hebb learning, Fuzzy c-means clustering, K-means clustering and Self-organizing maps(SOM).

Keywords:- background subtraction, foreground detection, feature extraction, pattern recognition, artificial intelligence.

I. INTRODUCTION

The Internet Of Things have successively advanced the communication between physical devices by collecting, exchanging, storing and processing data. Sufficient intelligence to automatically scale to the required timing of data warranting coupled with feasibility have been employed in various facets of Engineering. One such application is the Body Worn Camera(BWC), a video recording system which is mostly used by the defence sector for law enforcement. A smart redaction system enables for a quicker release compared to manually publishing. BWC differs from a CCTV as there are constraints on spatial area and conversation recording in a CCTV. A BWC can be lodged into the officer's suit, merging between buttons or with sunglasses for effective placement. Technology has facilitated the user interface with multiple options such as touch-screen, push buttons to record, playback on field, etc. The evidence is encrypted and managed online through a digital media storage platform or stored locally through a docking station.

BWC has proved to be very advantageous in law systems in their day to day life such as:
Transparency: The releasing of videos at the crime scene increases the trust of the public in the police. Officer protection and training: BWC protects police from false allegations and influences in a positive way. It also educates the experienced officers. Investigative: The records of the BWC can be substituted for the recollection of crime events for replaying and documentation purposes.
However, there are certain drawbacks prevalent to BWC that calls forth for further development. A camera only records in 2D and therefore accurate judgement regarding the distance cannot be inferred from the videos. A camera cannot be cited as the sole proof for a thorough investigation and sometimes requires one or more cameras to aid in the process. Also, due to the dearth of storage capacity and power, BWC remains unused in lot of domains. We suggest a BWC that is always on and automatically clicks pictures based on certain postures which is given as an input to the database. This survey aims at the different techniques that are incorporated for pattern recognition of images captured in a body worn camera using artificial intelligence techniques.
In this paper, we expound the use of a body worn camera predominantly in the defence sector. Exploration is done on the efficient methods for video frame acquisition. There are various noises present is eliminated through background subtraction methods and subsequently fed into the texture feature analyzer. We then employ the expertise of Artificial Intelligence approaches to classify the images, majorly under supervised and unsupervised learning techniques.

II. RELATED WORKS

A reaction of few cities in the United States to the idea of implementing BWC was mentioned in a report on body worn cameras by Eugene P. Ramirez. It cites the Rialto study. Rialto, was the first police department which did a study on the effects of the usage of BWC. It gives an overview of the Rialto study in which 2 groups of police were formed. The first were required to wear BWC during their shifts and the second were not. The use of BWC resulted in the following: Reduced the use-of-force incidents by 59% and reduced the citizen's complaints by 87.5%. The difference between CCTV and BWC that is mentioned in the Rialto study is given. CCTV is not audience specific and hence can work as a drawback because sometimes it invades privacy. Whereas, BWC comes only when the police is interacting and usually with suspects/offenders. Plus, mostly the consent of people is also obtained before recording on BWC's. On these lines, it also demonstrates the ACLU(American Civil Liberties Union) and PERF(Police Executive Research Forum) which state the positive and negative effects of BWC and a few recommendations that helps in successful deployment. The report finally suggests in favour of using BWC's along with mentioning that it should not be the sole factor considered for judgement.

An overview of how the use of BWC has increased in less than a year is presented in Existing and Ongoing Body Worn Camera Research for the Laura and John Arnold Foundation. It has transformed from receiving less support to currently being funded and is being used on a large scale. It states that, despite all this, only superficial knowledge is known about this technology. Research is limited to its effect on the police behaviour and very little study/work has been proved using scientific methods. Hence it suggests that studies should include both police and citizens and legal analysis should be done regarding this. Emphasis is to be given to privacy, redaction and storage of BWC data. It works on 4 phases and has produced results for the first:

Yet another paper explores a camera front end with Restricted Boltzmann Machine based Artificial Neural Network as the recognition and classification engine. An Ultra-low power, "Always-On" Camera Front-End for Posture Detection in Body Worn Cameras using Restricted Boltzmann Machines by Soham Jayesh Desai et. al, uses Weizmann Human Actions Silhouette database. The hardware implemented on Xilinx virtex 7 XC7VX485T which always-on sensing and recognition of 19.18uW. Through this the background extraction performed on ASIC has shown to consume around 27.88uW/ pixel. But this method does not aid for very deep networks as they consume tens to hundreds of thousands neural processing units. The cost is prohibitive in a mobile platform and they are not suitable for a mobile platform. They achieve up to 90% accuracy through Restricted Boltzmann Machine (RBM) based on Artificial Neural networks. MATLAB is used on the dataset and the weights are then transferred to the Xilinx compiler using Memory Initialization Files. To minimize the overall network power utilization at real time and accuracy performance constraints they choose 8 bits for data representation and 300 as the number of virtual NPC's. They obtained a processing time for classification of a single input frame 58.66us.They study the trade-off of power, timing work configuration selected for experiment is 256 feature inputs, 300 virtual NPC and 30 physical NPCs for hidden layer 1 and a final layer comprising of 10 NPC corresponding to 10 silhouettes actions. During design they used vector based estimation for the synthesized design. As a result, they spent less than 5uJ per frame and achieve an accuracy of 85% for a limited training set and shallow network while still maintaining the real time constraints.

Image filtering, reconstruction and compression is explained in Enemy Identification Technique in Defence Sector Using Digital Signal Processing and Digital Image Processing by K.S Mahalakshmi et al. The reconstruction is implemented by equalising the gray scale characteristics of the image, eliminating noises. The implementation is done in MATLAB tool. The paper mainly uses Fast Fourier Transformation techniques which is an inevitable image processing tool to decompose an image into its sine and cosine components. The Proposed method works best with JPEG and PNG format images as the compression, transformation and the reconstruction methods are easy and the storage requirements scale down. They propose a speedy reconstruction technique based only on software with no hardware requirements, the whole process terminating in about 35-40 to produce the output.

A Low-Power 65-nm ASIC Implementation of Background Subtraction by Harish Bhaskar et al. The subtraction is carried out by a low power implementation of mean-filter based background subtraction block in ASIC flow using 65nm CMOS process technology. The design is modelled in verilog HDL. The frames are assumed to be stored in memory an input to the circuit one pixel(byte) at every clock cycle. The max operating frequency is 800MHz which provides the ability to process up to 385 HD fps. The total power consumption as observed is 27.88 µW/pixel , which makes it suitable for low power applications. The image background and foreground is filtered using the mean-filtering process as the frame differencing is usually inaccurate in detecting the interior pixels of large and uniformly coloured moving object.

To improve the performance of the current methods of object recognition, a large dataset has to be trained like in Image Net classification by deep convoluted neural networks by Geoffrey E.Hinton et al. When Image Net was used in the LSVRC-2010 and ILSVRC-2012 competitions and achieved record breaking results. A highly optimized GPU implementation of 2D convolution and all the other operation inherent in training convolutional neural networks. They make use of Rectified Linear Units (ReLus), which work several times faster than their

equivalent with tanh units in the standard approach equation. The net is spread across two GPU's. The overall architecture consists of eight layers with weights. The first five are convoluted and the remaining three are fully connected. The output of the fully connected layer is fed to a 1000-way softmax which produces a distribution over the 1000 class labels. This network maximizes the multinomial logistic regression objective. The models were trained using stochastic gradient descent with a batch size of 128 examples, momentum of 0.9 and weight decay of 0.0005 which was a small amount but was substantial for the model to learn.

Real-time foreground-background segmentation using codebook model. A codebook algorithm is introduced by David Hardwood et al. The algorithm is aimed to sample the values over long times without assumption of parameters. Different backgrounds can be modelled by many codeword's. The algorithm has: a small background model to capture the moving/changing background, moving foreground images in the scene, allowance of multiple background layers through layered modelling and detection. Construction of the codebook algorithm: An initial algorithm for colour image and gray scale image is done by using training sequence, codebook for a certain codeword's, RGB vector and intensity values. This product is known as fat codebook. Next, the codeword's having foreground objects from true background are separated. For colour and brightness, the pattern of change of pixel values over time under light is observed. Based on this, a colour model is introduced to separate colour distortion and brightness distortion. The background subtraction algorithm is used for foreground detection also. Comparison is done on MOG and Kernel background modelling techniques with their method. The improvements are done by layered modelling/detection and adaptive codebook updating. The performance of several background subtraction algorithms is tested by PDR(perturbation detection analysis).

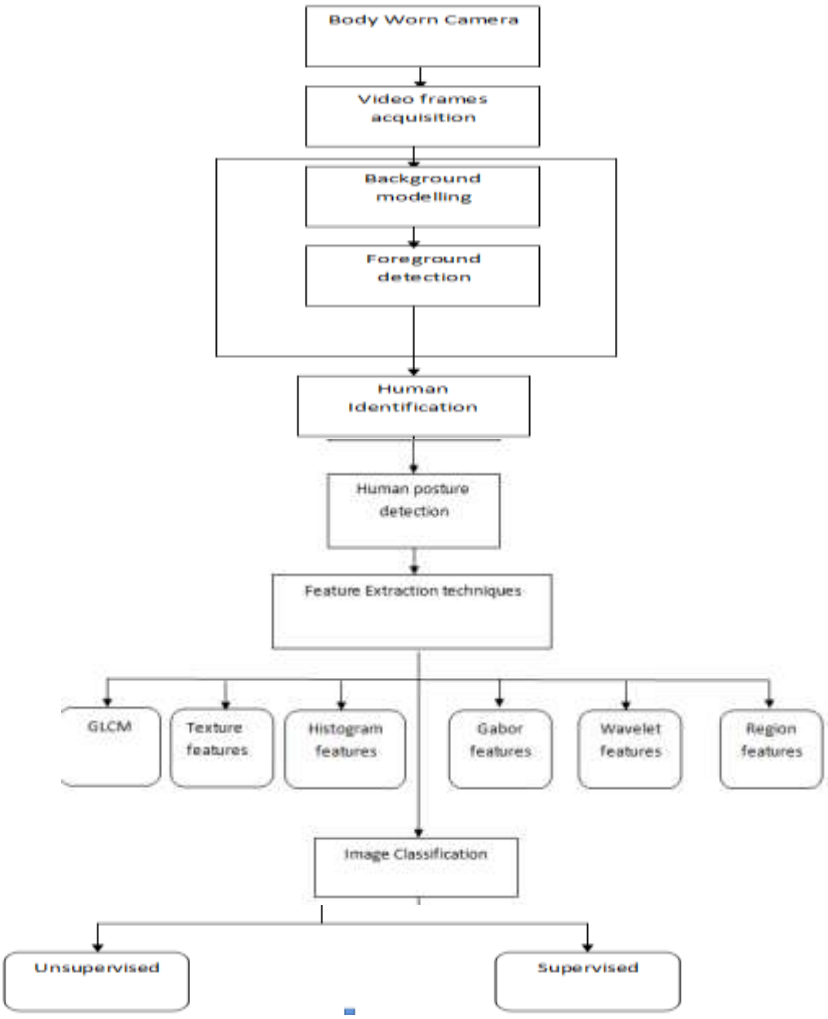
Human Silhouette Extraction on FPGAs for Infrared Night Vision Military Surveillance by Usman Zakir et al. An intelligent infrared video surveillance system is required to provide efficient analysis of the scene and detect human targets to assist the military in correct judgement. In the military, the FLIR surveillance is apt with it's features: portable, low-power and real-time. Several approaches to motion segmentation have been done and background subtraction is considered to be the best. However, it is sensitive in dynamic scenes. After the background is modelled, it is stored. The motion-based silhouette extraction is done by: Denoising the video control, background modelling, foreground estimation, silhouette extraction. So, a co-processor for infrared visual surveillance is developed to estimate the foreground images and extract what is necessary. It is coded and simulated using VHDL. With these, military night vision surveillance can detect human motion efficiently.

Pre-processing:

Pre-processing is done on images containing text/graphics. The resulting image after pre-processing is a binary image which contains only text. The various steps involved in this technique are:

- The image is made devoid of all extraneous noises.
- The watermarks or scenes in the background are removed
- Text and graphics are separated
- Characters are then separated from each other by using character segmentation
- Morphological processing is done to the characters that have been altered(parts of characters removed or added) by other pre-processing techniques.

III. SYSTEM ARCHITECTURE



- Body Worn Cameras (BWC): As mentioned in the introduction, **Body worn cameras** are a video surveillance system that is usually used by law enforcement to record videos and act as evidence at crime scenes, and is very useful in helping officer and public accountability. The implementation of an always-on, low power BWC is effective as is capable of making decision and starts to capture when the posture detected matches the training set.
- Video frame acquisition: The BWC's have different types of camera and various options to record the video, given audio and video feedback, then playback. Video is then uploaded from the docking station onto a local storage device such as a server or any web-based platform or sometimes in the field only. The video that is recorded in the BWC is in a standard and open format such that it can be replayed in any freely available software such as VLC without any processing and conversion required. It has standard resolution.
- Background Subtraction:

❖ Background modelling: It is an imperative step for object detection in many video processing systems. The image can be modelled using verilog HDL. The first four frames could be used to model and the method would be applied to the current frame and tested for threshold intensity, which can also be estimated using the histograms of the image and the result is stored in memory.

❖ Foreground detection: This mainly deals with detecting changes in a set background and therefore should be immune to temporal changes like lighting, repetitive movements, etc. After the background modelling, the image can be filtered using the mean-filtering process which accounts for an easy and effective methodology.

- Human identification: The aim of body worn cameras in defence sector to detect human actions tends towards aggression. The image that is filtered is identified to be containing a human silhouette.
- Posture detection: Once a human silhouette is identified, the posture detection decides whether there is an intention of aggression and harm. To achieve this, the model is trained with Weizmann human actions database which has concentrated sets of images for aggression.

- Feature extraction techniques:

❖ GLCM: Known as the gray level co-occurrence matrix, GLCM is a statistical method of texture analysis after noting the spatial relationship of the pixels. The image can be classified into a number of orders where GCLM works mainly in the second order texture feature extraction. The various characteristics / features of GCLM are: energy, entropy, contrast, homogeneity, correlation, shade and prominence.

❖ Texture feature: The feature extraction involves extracting the features from high resolution remote sensing images. It results as a complementary source of data for those applications in which the spectral information is not sufficient for identification. Few of the texture feature extraction methods include: statistical methods, filter techniques, and the latest wavelet decomposition.

❖ Histogram features: The Scale-invariant feature transform(SIFT) detects the corners and edges of the image. In the resulting image, it finds the points or Regions of interest(ROI). From each of these, a histogram is extracted from which has a particular corner/edge orientation. These histograms are then quantized into a smaller number of groups by K-means(clustering method).

❖ Gabor features: A set of Gabor filters (a sinusoidal signal) are used which have different frequencies and orientation. Using these, a host of features can be extracted such as edge and object detection, image representation and coding, colour or pattern gradient, etc.

❖ Wavelet features: The wavelet feature does: compact most of the signal's energy into a few transformation coefficients, capture and represent effectively low frequency components (such as image backgrounds) as well as high frequency transients (such as image edges),the variable resolution decomposition with almost uncorrelated coefficients, progressive transmission, which facilitates the reception of an image at different qualities.

❖ Region features: Region extraction is an image segmentation technique that categorizes the pixels in an image as belonging to one of many discrete regions. The process consists of scanning the new regions that are met and merges old regions when they are connected on a lower row. The image is scanned and each pixel is labelled with an identifier which specifies the region to which it belongs.

- **Image Classification**

Once the feature extraction terminates, the image can be classified by certain pattern recognition methods.

There are two main methods for classification:

- Supervised method
- Unsupervised method.

Supervised method: There is a definite and desired output in connection with the labelled information given as the training set to the agent. It further contains these algorithms as we have mentioned in the architecture diagram.

a. Artificial Neural Network(ANN):Problems comprising of noise, imperfection, nonlinearities and so forth can be treated by ANN. Supervision teaches the neural network to construct a system from the fed data. It consists of two distinct methods: Texture feature extraction and training. Texture feature extraction is implemented using the GLCM method as explained above. The features are extracted from n x n primary matrix in directions starting from degree 0 with a successive increase of 45 degrees to make them direction invariant .Training consists of a three layered neural network: Input, hidden and output. The input layer is made up of values of the data which is fed as inputs to the next layer of neurons. Following which, there might be multiple hidden layers. The output layer consists of one node for each class. A single transaction through the network attributes a value to an output and the record attributed to the class node with the highest value.

b. **Decision trees:** This method is optimal for large sets of data. Here the attributes are split in such a way that at each branch, the partitions are as pure as possible. A general decision tree algorithm: Input data is fed in the form of data partition which is a set of training tuples, attribute list and an attribute selection method. The method creates a node. A leaf node is created if the tuples are all of the same classes. If the attribute list is empty, then a leaf node labelled with the majority class, is returned. This is also called as majority voting after which the attribute selection method is applied to find the best splitting criterion. For discrete value splitting attributes, multi-way splitting is allowed. Then the splitting attribute is removed. For each outcome of the splitting criterion, the tuples are partitioned and a sub-tree is grown. In the end a node is returned.

c. **Minimum distance:** As specified by the unique parameter, about 2^7 unique class segments are utilized to classify data. The mean vector in each class is used by ignoring other methods such as covariance matrix, standard deviation, etc. Following these, a theme map is generated. A unique gray level is encoded with each class. To facilitate better understanding, a table containing information about the specified colours

should be presented when it is given to the display as the gray level is employed when the unique class signature is created.

d. **Bayesian network:** This method relies on the theory developed by Thomas Bayes. Based on the probability of each cause and the conditional probability of the outcome of each cause, the conditional probability of a set of possible outcome is generated. The Bayesian Net classifier trains the data .Various search algorithms and quality measures are utilized by the bayesian net learning. There is a base class that distributes data structures like network structure, conditional probabilities, etc. An estimation method using a simple estimator which is computed directly from the data. A genetic search is used to improve the net structure. This genetic search enables crossover and mutations to obtain off-springs. The best structure found is returned.

e. Support Vector Machines(SVM):They are mainly used for classification and regression analysis. Points in space are mapped such that the members of distinct entities stand out with an evident space. Then, based on the side of the space which they fall, they are attributed accordingly. The missing values are supplemented and the nominal attributes are transformed into binary ones. Normalization is done by default. The coefficients in the output are derived from the normalized data. Multiclass problems are classified pair wise. The choice that matches the logistic regression model to the output of SVM is selected. A hyper plane is a set of points in high dimensional space whose dot product with a vector is a constant. The vectors in hyper planes could be in linear combination with the parameters of the images of feature vectors. The sum of kernels can be used to measure the relativity of each sample point to the data point, emanating from one or the other set to be discriminated.

Unsupervised Method: This method of pattern recognition identifies similar patterns within a data set. They generally assume that there is an underlying pattern with the data. They see patterns in random data and hence their output must be continuously validated statistically and scientifically.

a. Hebb learning: This learning method is based on Donald Hebb's findings. The association is done with the brain cells, the neurons. Simply stated, if two neurons are present on either side of a synapse/connection and are activated simultaneously, then the strength is increased. This theory is often complemented by :If the neurons present on the either side of a synapse/connection are activated asynchronously, then the strength decreases. The hebb algorithm for pattern association : All the units of the image are assigned weights and are all initiated to 0.The input and output units are allotted successive integers for the current input and current target output .The weights are then modified accordingly where the new weight is updated. Then the activation of the output units is set.

b. Deep Learning Methods: This is used in the field of neural networks. The input to the network consists of pixels They carry out tasks that are difficult for computers (such as pattern recognition). It addresses the problems that arise when back propagation algorithms to deep networks that have many layers. This is applied in practice through Boltzmann machines or auto encoders in each hidden layer as building blocks in forming deep neural networks.

c. Fuzzy c- means clustering: Traditionally in hard clustering, a data member was not permitted to be a part of another cluster. But through Fuzzy c-means clustering, a membership with multiple clusters is possible. Each data member is assigned a weight based on the distance from a cluster centre where the weightage decreases as the distance increases from the cluster centre. The algorithm begins by selecting a number of clusters. Then coefficients are assigned for being a part of a cluster. The method is repeated until the successive change between two iterations does not overshoot the allotted sensitivity level. This is called converging following which the centroid and the member coefficient is calculated. In image processing, the image features as fuzzy sets. The entire procedure summarizes into: Image fuzzification, the membership values are modified and if mandatory, image defuzzification.

d. K-means clustering: Vector quantization that arises from signal processing and used in data mining. It partitions n observations into k clusters in which each observation belongs to a cluster which has the nearest mean and acts as a prototype of the clusters. This also partitions the data space into Voronoi cells.

e. Self-organizing maps(SOM): SOM aims to organise data within the image distinctly in the given input space. Like all neural networks, this too consists of training and mapping. Classification is done by the following steps: The weight vectors are initialized with the chosen random values. A sample input vector is selected . The neurons that is a closest match to the selected vector is found. Finally it is updated with the weight update equation followed which the steps are repeated from selecting a sample vector until feature map stops changing.

Artificial Intelligence Applications

The various AI techniques are Neural Network, Fuzzy Logic, Genetic Algorithm and Hybrid AI. The different **Areas Of AI Are:** language understanding, problem solving, learning and adapting systems, perception (pattern recognition and scene analysis), modelling, robots and games.

- AI in design of PSS(Power System Stabilizer) : The techniques generally used are artificial neural networks, fuzzy logic.
- AI in network intrusion detection: Intrusion detection systems (IDS) is monitoring the events that occur in the network and detection of intrusions. Here again, the techniques used are Artificial neural networks (ANN) and Fuzzy inference systems in IDS.
- AI in the area of medicine: Fuzzy expert systems in medicine(data handling method that allows ambiguity and is suitable for medical applications), Evolutionary computation in medicine(computer techniques that imitates mechanism of natural selection in solving real world problems)
- AI in hospital inpatient care: Helps in medical image classification(ANN in diagnostic science such as endoscopic image, MRI brain tumour analysis)
- AI in accounting databases: AI builds intelligence into database to help users(helps in sorting large data and provides assistance in searching and evaluating data within the time constraints).

IV. CONCLUSIONS

Artificial intelligence approaches come handy in handling uncertainties, error prone data, problems associated with imperfections, nonlinearities, so on and so forth. This paper aims to boost the technology offered by the Body Worn Camera. In future, we seek to implement the various methods proposed above and attempt a comparison between them for efficiency and suitability under multiple constraints. This would serve as the base paper, rooting on which the best methods pertaining to each category of data would be divulged, hence reinforcing the arms for identification of criminals, catalysing security.

ACKNOWLEDGMENT

We would like to take this opportunity to express our sincere gratitude towards Dr.J.Preethi, M.E.,Ph.D., Assistant Professor of the Department of Computer Science and Engineering at Anna University, Regional Centre, Coimbatore, for her guidance and constant support throughout the course of the paper. Her perceptive criticism fuelled our motive to work better and achieve a refined draft of the paper. Her core knowledge in artificial intelligence helped to edit better and kept us from detour.

REFERENCES

[1]. 1Jitendra Singh Sengar, 2 Niresh Sharm "Design a Neural Network Based on Hebbian Learning and ART" 1,2Dept. of CSE, RKDF IST, Bhopal, MP, India.

[2]. Madhu Bala"Comparative Analysis on Scene Image Classification using Selected Hybrid Features " Myneni Professor, Dept. of CSE, MIST, Hyderabad-5, A.P, India. M.Seetha Professor, Dept. of CSE, GNITS, Hyderabad, A.P.

[3]. A Review Priyanka Sharma Manavjeet Kaur Department of Computer Science and Engineering Department of Computer Science and Engineering PEC University of Technology, Chandigarh, India PEC University of Technology, Chandigarh, International Journal of Advanced Research in Computer Science and Software Engineering Research Paper Available online at: www.ijarcsse.com Classification in Pattern Recognition India Volume 3, Issue 4, April 2013 ISSN: 2277 128X.

[4]. Kyungnam Kim1 , Thanarat H. Chalidabhongse2 , David Harwood1 , Larry Davis1} "BACKGROUND MODELING AND SUBTRACTION BY CODEBOOK CONSTRUCTION" 1Computer Vision Lab, University of Maryland, College Park, MD 20742, USA 2Faculty of Information Technology, King Mongkut's Institute of Technology, Thailand 1{knkim|harwood|lsd@umiacs.umd.edu}, 2{thanarat@it.kmitl.ac.th

[5]. A Tutorial of Wavelet for Pattern Recognition r99942126 Guan-Chen Pan 潘冠臣 E-mail: r99942126@ntu.edu.tw Graduate Institute of Communication Engineering National Taiwan University, Taipei, Taiwan, ROC

[6]. P. Mohanaiah* , P. Sathyanarayana** , L. GuruKumar*** * "Image Texture Feature Extraction Using GLCM Approach" Professor, Dept. of E.C.E, N.B.K.R.IST, Vidyanagar, Nellore, India ** Professor, Dept. of E.C.E, S.V University Tirupati, India *** Asst.Professor, Dept. of E.C.E, N.B.K.R.IST, Vidyanagar, Nellore, India.

[7]. Namratha M 1 , Prajwala T R "A Comprehensive Overview of Clustering Algorithms in Pattern Recognition" 2 1, 2(Dept. of information science, PESIT/visvesvaraya technological university, India)

[8]. Avneet Pannu, "Artificial Intelligence and its Application in Different Areas" M. Tech Student Department of Computer Science & Engineering DAV Institute of Engineering and Technology, Jalandhar India.