RESEARCH ARTICLE OPEN ACCESS

Development of Algorithm for Object Detection & Tracking Using RGB Model

Vivek Tiwari [1], Dr. Vivek Sharma [2]
Scholar [1], Professor [2]
Department of Computer Science and Engineering
RGPV/TIT, Bhopal M.P - India

ABSTRACT

This research work introduces a method of using color thresholds to identify two-dimensional images in MATLAB using the RGB Color model to recognize the Color preferred by the user in the picture. Methodologies including image color detection convert a 3-D RGB Image into a Gray-scale Image, at that point subtract the two pictures to obtain a 2-D black-and-white picture, filtering the noise picture elements using a median filter, detecting with a connected component mark digital pictures in the connected area and utilize the bounding box and its properties to calculate the metric for every marking area. In addition, the shade of the picture element is identified by examining the RGB value of every picture element present in the picture. Color Detection algorithm is executed utilizing the MATLAB Picture handling Tool kit .We also want to find features to classify the foreground moving objects, an easy task 1 above the floor. From this video troublesome parts have been selected to test the changes done to the algorithms and program code.

Keywords:- MATLAB, picture handling tool kit, shading location, RGB picture, picture division, Image filtering, bounding box. High Definition (HD)..

I. INTRODUCTION

The world has witnessed a surging demand for videos distributed through the Internet. In many cases, instead of playing the video after downloading the complete file, the user can watch the video in real time via a data link connected to the server. We call this video streaming. Assuming the user plays the video at the same rate as it downloads the video, the simplest way of streaming from a server.

For video content providers who do not have a budget to expand the physical uplink bandwidth, a natural question to ask is: Does there exist a way to serve a large number of peers who are trying to stream the same video, given a limited upload bandwidth from the server?

Video surveillance data processing is currently accomplished by techniques such as parallel computing and distributed computing to reduce costs. Such techniques provide performance enhancement and reduction in cost; yet suffer from limitations in resources, complex programming, scalable storage and limited support of fault tolerance. When considering these challenges within the current infrastructure, a data processing framework that is simple and automatically handles task scheduling, distribution and storage of data, load balancing, and machine failure is necessary in order to allow users to focus solely upon creating scalable applications.

The central idea for a Peer-to-Peer (P2P) video streaming system is to fully utilize every user's upload bandwidth and storage space to help the server distribute the content. Instead of download only from the server, each user now is allowed to contact other peers and download parts of the file. Hence the total upload service requirement is now shared among the group of peers, rather than by the server alone.

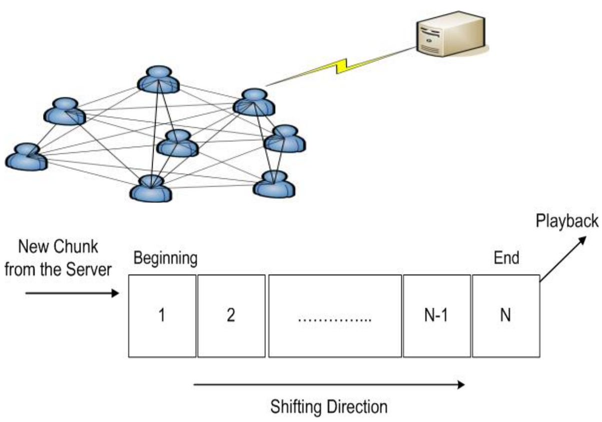


Figure 1.1: Structure of a Streaming Network

Consider a network with M peers trying to stream a live video from a single server. For simplicity of analysis, assume the system progresses on a discrete time basis (t = 1; 2; 3; :::), video contents are distributed in the form of data packets called chunks, and the video is played at the rate of one chunk per time slot. At the beginning of each time slot, the server selects a peer uniformly at random from the network and uploads the most recent chunk. Due to the limited upload bandwidth at the server, each peer employs an N-chunk buffer to cache the video stream, where location 1 denotes the most recent chunk that is currently being uploaded by the server, and location N is where the video is currently being played. During each time slot, every peer is allowed to contact another peer randomly selected from the rest of the network, which we call target, and download at most one chunk of content that is available in target's buffer locations 2 to N - 1. At the end of the time slot, the chunk at buffer location N, if it exists, is played, and all other chunks in the buffer are shifted by one location towards the end of the buffer.

HOW TO MANAGE VIDEO PROCESSING SYSTEM

Large-scale distributed systems are required for video surveillance systems (VSS) in order to analyses large quantities of recorded video data which is a computing intensive activity. It is important to consider scalability as a factor for future video surveillance systems []. Existing solutions require demand in resources, which are unsuitable for future increased demands for video data. IBMs system, IBM Smart Surveillance System, (IBMSSS) [], deploys a combination of database partitioning and web application server clustering that allows scalability. However, such solutions that attempt to resolve the scalability issue are expensive and increase the cost of hardware and overall investment expenses.

II. MOTIVATION

A literature review was undertaken in order to understand the current regulations and guidelines behind establishing a reliable, legal and trustworthy, cloud based video surveillance system. The requirements of a legally acceptable video forensic system are discussed and current security and privacy challenges of cloud based computing systems are studied in order to recommend the design of a secure and reliable cloud-based video forensic system. The research focuses only on the performance of distributed video analytic applications using a cloud-based Hadoop platform after carefully considering the proposed recommendations and observations.

Preliminary experimental results conducted within the context of the research presented in this thesis indicates that a cloud-based Hadoop platform will be successful in speeding up video analytic processing and distribute computing of computer vision algorithms in a cluster of machines. It was also observed that the execution time of applications is determined mainly by both the size of the load, as well as cluster sizing. Hadoop Map-Reduce architecture is dependent upon the type of application, as well as upon hardware performance/configuration[16]. However there was no existing research in attempting to model Hadoop performance within a distributed computing framework serving data and processing intensive applications.

Identifying the above research gap, the behavior and performance of a video analytic application running in a virtualized cluster is first studied in this thesis, which highlights the most significant factors that intense the execution time of an application. Based on these factors an experimental study was conducted in order to develop a prediction model for the application. This was undertaken by comparing different machine learning algorithms based on the prediction accuracies that reveal that decision-based models outperform linear regression models, whilst the Ensemble Bagging models outperform standard single-based classifiers. This research fall an existing gap in research relating to video analytic related comprehensive performance predictions. Current research maintains a focus upon different types of applications that are limited to using standard learning algorithms, such as Simple Linear Regression, SVM and Multilayer Perception (MLP).

III. LITREATURE SURVEY

Colour quality is an important consideration, yet there are not currently any well-accepted models of perceptual quality prediction of colour images. However, pretty good results and some improvement relative to ‘ ‘ luminance only’ ’ processing can be obtained by applying standard single-channel QA models to one or more chromatic channels, then combining these in various ways. However, progress remains to be made on this problem, given the complexities of colour perception and the lack of current models of colour distortion perception. Assessing the quality of stereoscopic (3-D) images is also a topic of pressing interest. The main problem is geometry and visual comfort, which is difficult to ensure without a large Hollywood budget (and even then!). However, our topic here is the perception of distortions, and the role of depth on distortion perception remains quite murky, and like colour, no entirely successful method of stereo image quality assessment has been found. While several approaches have been proposed for 3-D stereoscopic QA 2-D quality models applied to stereo pairs commonly perform as well as, or better than, 3-DQA models that utilize depth or computed disparity maps. in a scene to portray objects and figures more prominently in paintings a style popularized in Renaissance called Chiaroscuro. In this paper, we propose the first computational model named LuCo to capture and quantify this effect, introducing a measure of lighting surprise. The Bayesian surprise value is first calculated based on a set of local observations of luminance patches. Precision matrices of a GMRF model for the prior and posterior distributions are estimated via sparse graph learning. Finally, a histogram of the acquired surprise map is constructed, and the computed parameter of the histogram is deemed the LuCo score. A large LuCo score thus means that luminance surprises are either very big or very small in a painting, reflecting the artist' s intention to accentuate lighting contrast. Experimental results verify the effectiveness of our LuCo metric when compared to a gradient based method and saliency-based methods. In particular, our computed LuCo scores reflect different usages of lighting contrast in Chiaroscuro and Impressionism paintings, and capture the luminance contrast changes throughout Rembrandt' s life.

Such measures as LuCo open possibilities towards preservation of artistic intentions during media delivery [15].

SOFTWARE USED :

MATLAB (Matrix Laboratory) is a numerical computing upbringing and 4&5 generation invention encoding language. Launched by Math-works, MATLAB allow matrix manipulations, intrigues of functions and data, execution of algorithms, construction of user interfaces, and interfacing with programs written in other languages, including C, C++, Java, and FORTRAN.

In the problem of video processing we are using MATLAB R2013a. For this specific purpose we are using communication tool in MATLAB. The communication Toolbox is a collection of computation functions and simulation blocks for research, development, system design, analysis & simulation in the video processing area. The toolbox is designed to be suitable for experts and beginners in the communication field. MATLAB R2013a version is used for simulation purpose.

IV.OVERVIEW OF SYSTEM

Human sensory processing capabilities are limited in the domain of sound, touch, and smell but well developed in processing visual patterns. The means of human visual information capture are the eyes, and studies have shown that they have limited range of visual perception, but good at discriminating features. Man is not unique in processing sensory information since other animals such as whales have well developed sound processing capabilities, and rely on them for food and protection. For example, ants communicate using smell from pheromones deposited on the ground wherever they visit to assist the colony in search of food.

Table1.1 is an evaluation of surveillance activities based predominantly on humans against requirements of performance (accuracy, timeliness and reliability), cost-effectiveness and user-friendliness. The main problems with intelligence gathering centered on humans are: the slow means of information transfer from gatherers to centre of intelligence, slow means of transfer of decisions and commands from centre of intelligence to action implementers, and low volume of information transferred per trip. Visual communication using manual processing (rely predominantly on humans) is relatively slow, expensive and inefficient especially when visual information is to be gathered over large area of coverage.

A solution to the high cost of gathering information over large areas is the use of image acquisition devices (cameras, infra red and thermal imaging device). Closed circuit television (CCTV) cameras in particular provides a cost-effective means of acquiring images on a continuous basis, and over large areas using multiple cameras. In response to increasing volume of visual data continuously being acquired storage devices are used for archiving and playback of video streams.

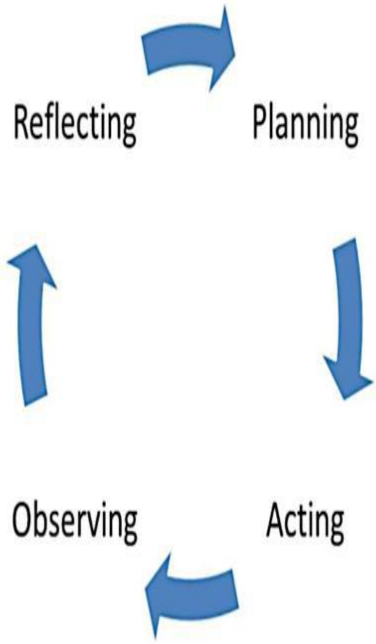
The following are possible approaches to solving these problems: the problem of continuous mode image acquisition, analysis, and instantaneous decision making capabilities on a large scale deployment scenario could be solved using, computer based systems with distributed processing, centralized/distributed monitoring and control of operations and rapid response to event in progress. The processing system must be characterized by scalable computer processing power to match required processing power, and scalable processing techniques (parallel/distributed algorithms for robust content analysis); and real-time processing capability to meet application requirements.

VISUAL SCENE ANALYSIS

Typical visual scene analysis algorithms involve the following sequence of tasks: pre processing, object detection, object tracking, and anomalous behavior detection. Pre processing typically involves frame format inter conversion, noise removal, decompression, and object enhancement. Object detection typically involves scene modeling, candidate object localization, analysis/synthesis of candidate objects, classification and detection, and anomalous behavior analysis. Object localization typically involves identifying locations of likely objects. For a given object location object analysis or synthesis technique is applied to identify its features or to model the object. When several objects are of interest in a scene then one object class must be differentiated from another object class, hence objects must be classified. Also in detecting single objects, background objects would have to be differentiated from the object of interest. Object classification may be part of an object detection task since a particular object in a class might have to be identified from among other objects not in the same class. Detection typically follows classification and involves evaluation of confidence level after classification or some validation test. The output from the detector is typically the location, and the class of the candidate objects. Object tracking involves establishing correspondence between the same object in different frames. Anomalous behavior detection involves defining atypical behavior as a sequence of discrete events. Continuous mode visual scene analysis operating twenty-four hours a day is faced with several challenges including the following:

Analysis complexity: Increasing analysis complexity typically arises in complex scenes involving illumination changes, scene clutter, scale changes, camera motion, and low object background contrast. For instance changes in scale brought about by perspective projection due to object moving away from a stationary camera might make a feature-based detection technique fail due to difficulty in differentiating object features from noise at very low object resolution. Similarly, the choice of object models on which object analysis and synthesis depends on has direct effect on complexity. For instance 3-D models of humans and its associated motion models are computationally demanding, although the accuracy is better compared to 2-D models [26, 27]. The choice of algorithms and the assumptions on which it is based also has direct effect on analysis complexity. With 2-D motion models, an assumption of smoothness of motion or changes in illumination is used in motion tracking, or optical flow to reduce analysis and computational complexity. Similarly in tracking, multiple hypotheses tracking with exponential search complexity could be avoided by excluding certain incompatible events from occurring simultaneously.

Accuracy: The accuracy of object detection and tracking measures how often the system makes correct and incorrect detection and tracking decisions and the confidence levels associated with this decision process. The accuracy of detection and tracking objects in visual scene is dependent on whether objects exist in isolation or part of a group, besides scene complexity factors. As a general observation, objects in a group tend to occlude features of each other. For example two humans moving together as a group might result in features of the person.
example, holding weekly meetings or information sessions with employees c) Observing the impact of actions, for example, speaking to staff to assess whether or not they felt more informed and/or empowered by information meetings d) Reflecting on the impact of actions, for example, what have I learned about myself? How do I know if staff feels more informed? What should I do differently next time? Upon reflection, the cycle begins again using a revised plan. That contended that this cycle could not only improve practice by taking action, but that these improvements were made in a democratic manner which took into account the views of staff that were affected by decisions, thus improving the situation for all parties involved.



Object detection
Object detection is identifying objects of interest in the video sequence. Object detection can be done by various techniques such as frame differencing, optical flow and background subtraction.

- **Object Classification**

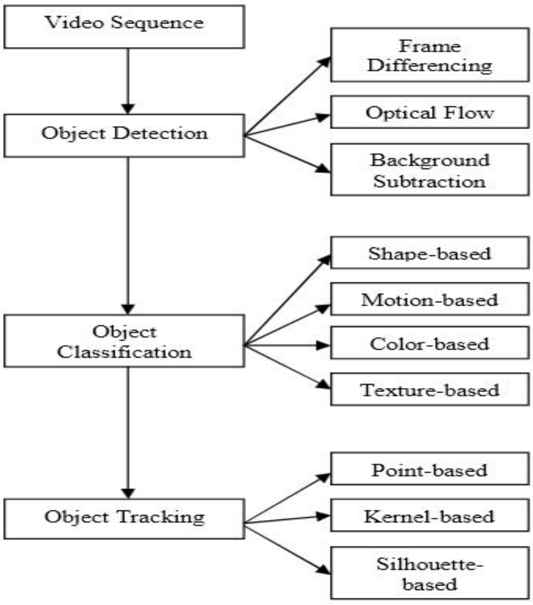
Object can be classified as vehicles, birds, floating clouds, swaying tree and other
moving objects. The approaches to classify the objects are e.g. shape-based classification,
motion-based classification, color based classification and texture based
classification.

- **Object Tracking**

Tracking can be defined as the problem of approximating the path of an object in the image plane as it moves around a scene. The approaches to track the objects are
point tracking, kernel tracking and silhouette.

- Noise in an image.
- Difficult object motion.

- Imperfect and entire object occlusions.
- Complex objects structures.



V. METHODOLOGY

In our work, we have tried to control mouse cursor movement and click events using a camera based on color detection technique. Here real time video has been captured using a Web-Camera. The user wears colored tapes to provide information to the system. Individual frames of the video are separately processed. The processing techniques involve an image subtraction algorithm to detect colors. Once the colors are detected the system performs various operations to track the cursor and performs control actions, the details of which are provided below. No additional hardware is required by the system other than the standard webcam which is provided in every laptop computer. System Approach Capturing real time video using Web-Camera.

- Processing the individual image frame.
- Flipping of each image frame.
- Conversion of each frame to a grey scale image.
- Color detection and extraction of the different colors
- (RGB) from flipped gray scale image.

COLOR DETECTION

Color detection involves detecting the color pixels of the tapes bearing on the finger tip in an image. It is a fundamental step.
A wide range of image processing applications such as face detection, hand tracking and hand gesture recognition [8][2].. The result would be a grayscale image (back projected image), where the intensity indicates the likelihood that the pixel is a color tape pixel. This method is adaptive since the histogram model is obtained from the users color caps, under the preset lighting condition.

1. Hand Contour Extraction

The OpenCV [1] function cvFindContours() use es an order finding edge detection method to find the contours in the image. In the contour extraction process, we are interested in extracting the hand contour so that shape analysis can be done on it to determine the hand gesture.

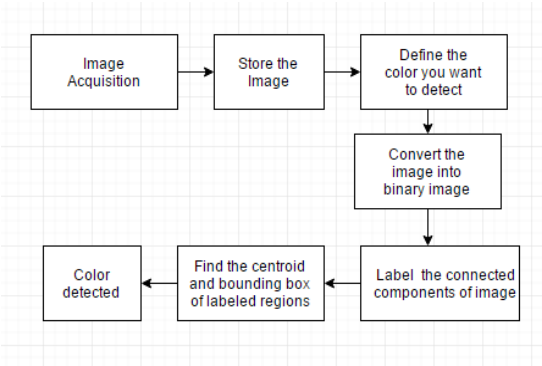
2. Gesture Recognition

The gesture recognition method used in the proposed design is a combination of two methods, the method proposed by Yeo and method proposed by Balazs. The algorithm for the proposed gesture recognition method is described in the flow chart that cab be shown below. It can be seen that the convexity defects for the hand contour must firstly be calculated. The convexity defects for the hand contour was calculated using the OpenCV inbuilt function " cv Convexity Defects" . The parameters of the convexity defects (start point, end point and depth point) are stored in a sequence of arrays. After the convexity defects are obtained, there are two main steps for gesture recognition:

Algorithm for Color Detection in Images

Step 1: Read an Input Image.
Step 2: Creating Red, green and blue Color bands of the Image.
Step 3: Compute and plot the Red, Green, Blue Color band Histogram.
Step 4 :Convert Input Image to gray-scale image.
Step 5:Track BLUE entity in real-time, we have to subtract the BLUE segment from the gray scale picture to remove the red segments in the picture..
Step 6: Apply Median-filter to remove noise.
Step 7:Transform the gray-scale picture into a black and white Picture.
Step 8:Extract every pixels less than 300px.
Step 9:Select all the associated segments in the picture.
Step 10:Find Centroid, Area, Bounding Box using Region props.
Step 10.1: Apply rectangular box for each Blue color in Image.
Step 11: Stop

Flow Chart Of Algorithm



VI. SIMULATION AND EXPERIMENTAL RESULTS

Image data is characteristically of an elevated bandwidth. Distinctive raw PAL TV signal determinations require 20 MB/sec used for broadcast. A single Digital movie format frame is 7MB large. The mission of image compression is to decrease the amount of data used to store up an image exclusive of objectionable degradation in the perceived quality of the image. The term objectionable depends of course on the use of the image data. Sometimes fast/real time transmission is more important than image quality e.g. video over wireless. Other times quality is paramount as in Digital Cinema and DTV. To make matters more complicated the same image shown in different formats looks quite different e.g. while single convert on or after one DTV system to a further, degradations in the conversion do not demonstrate up on television sets, but if the same image is used for creating a Digital Cinema production [film format ads created from TV ads] then the degradations are very noticeable The numerical values used for these intensities are usually chosen such that equal increments in value result in approximately equal apparent increases in brightness.

Our video colorization method is tested in two ways. One is to use frames randomly selected from some standard video sequences as references to color their next frames in the sequences. Another way is to use one selected frame as a reference and colorize the following few frames. PSNR is computed for objective quality evaluation by our colorization method.

1. RESULTS- FRAME BY FRAME

Silhouette tracking

This method is occasionally called *region tracking*, or *region-based tracking*.

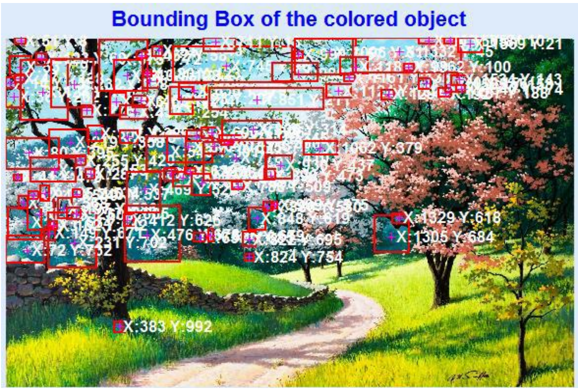
Kernel Tracking:-The *kernel tracking approach* is based on computing the motion of an object represented by a primitive object region. By using a motion model, i.e. computing the motion of an object from one frame to another.

Deterministic method-The development of future states in a system that does not involve any randomness is called a deterministic system. This means that using a given start condition or initial state in a deterministic model will always produce the same output.

IMPLEMENTATION:

Read the input image

Recognizing the boundaries of an object / Bounding box of the colored object-3D input images were converted to 2D arrays. Now select one of the random pixels in the background of the object and move it.



VI. CONCLUSION & FUTURE SCOPE

a. Conclusion

Unfortunately, these error measures don't align well with the Human perception of pictures. It' s an honest rule of thumb but that in comparison pictures victimization these measures, massive variations in objective measurements do tend to imply similar variations in human perception of the photographs. But once the variations between an equivalent objectives live on many pictures square measure little then it's not the case that the variations would imply similar sensory activity evaluations. An illustration throughout the lecture can justify this additional clearly. Victimization this technique, human subjects square measure asked to rank ascertained (processed) pictures in terms of this five purpose scale. Collating results from such experiments provides some quantitative measure of the subjective perception of the photographs.

b. Future scope-

Create and illumination variations still stay one of the largest challenges for face recognition. Most of the approaches which may handle non-canonical facial create need manual labeling of many landmark locations within the face that is AN impediment to realizing a completely automatic face recognition system. Another issue that must be addressed is strength to low resolution pictures. Common place face recognition ways work well if the photographs or videos square measure of high quality, however sometimes fail once enough resolution isn't accessible. Apart from these variations, lots of labor is needed to handle the opposite sources of variations like expressions, make-up, aging, etc.

Face recognition from video sequences may be a comparatively new analysis space as compared to still image based mostly face recognition and there are several unaddressed problems. The massive quantity of knowledge that's being collected from police investigation cameras requires economical ways for exploit, pre-processing and analyzing these video streams. Also, much work needs to be done to optimally use the dynamic info gift within the video sequence for recognition.

The brightness level and structure options are computed by intensity and LBP data, severally, for visual quality prediction of pictorial regions Use a dynamic game to formulate the behavior of crowd sources Mirrors with non-central viewpoint square measure used due to their specific properties. On the opposite hand, since they need non-central viewpoint, specific demands to style ways for proper Image formation or 3D purpose determination would be needed.

REFERENCE

[1] P.Sudharshan Duth,M.Mary Deepa, " Color detection in RGB – modeled images using MATLAB" , *IJET 7(2.31)(2018)29-33*.

[2] Wyszecki G & Stiles WS, *Color science*, New York: Wiley, (1982).

[3] Brownrigg DRK, " The weighted median filter" , *Communications of the ACM*, Vol.27, No.8,(1984), pp.807-818.

[4] Kittler J & Illingworth J, " Minimum error thresholding" , *Pattern recognition*, Vol.19, No.1, (1986), pp.41-47.

[5] Gonzalez RC & woods RE, Digital Image Processing, *Addison-Wesely Publishing Company*, (1992).

[6] Pal NR. & Bhandari D, " On object background classification" , *International journal of systems science*, Vol.23, No.11, (1992), pp.1903-1920

.
[7] Vincent L, “ Morphological grayscale reconstruction in image analysis: applications and efficient algorithms” , *IEEE transactions on image processing*, Vol.2, No.2, (1993), pp.176-201

.
[8] Solihin, Y & Leedham CG, “ The multi-stage approach to grey-scale image thresholding for specific applications” , *Nanyang Technological University, School of Computer Engineering, Nanyang Avenue, Republic of Singapore*, (2000)

. [9] Berns, RS & Reiman DM, “ Color managing the third edition of Billmeyer and Saltzman's Principles of Color Technology” , *Color Research & Application*, Vol.27, No.5,(2002), pp.360-373.

[10] Gonzalez RC, Woods RE & Eddins S.L, *Digital Image Processing Using MATLAB*, Pearson Education, Inc, (2004).

[11] McAndrew A, “ An introduction to digital image processing with matlab notes for scm2511 image processing” , *School of Computer Science and Mathematics, Victoria University of Technology*, (2004).
International Journal of Engineering & Technology 33

[12] Sezgin M & Sankur B, “ Survey over image thresholding techniques and quantitative performance evaluation” , *Journal of Electronic imaging*, Vol.13, No.1,(2004), pp.146-166.

[13] Abadpour A, Color Image Processing Using Principal, (2005).

[14] Senthamaraikannan D, Shriram S & William J, “ Real time color recognition” , *International Journal of Innovative Research In Electrical, Electronics, Instrumentation And Control Engineering*, Vol.2, No.3,(2014).

[15] Rani NS, Neethu OP & Ponnath N, “ Automatic Vehicle Tracking System Based on Fixed Thresholding and Histogram Based Edge Processing” , *International Journal of Electrical and Computer Engineering*, Vol.5, No.4,(2015).

[16] Rani NS, Verma SK & Joseph A, “ A Zone Based Approach for Classification and Recognition Of Telugu Handwritten Characters” , *International Journal of Electrical and Computer Engineering*, Vol.6, No.4,(2016).