ISSN: 2278 – 7798
International Journal of Science, Engineering and Technology Research (IJSETR)
Volume 5, Issue 9, September 2016

Vision Based Door Control System with Error Reduction

Swe Swe Win, Wut Yi Win

***Abstract*— This paper presented Automatic Door Control System Based on Face Tracking. The aim of this research is to develop and implement an automatic door control system without the need of manually opening and closing actions. This system will avoid the need to control by persons with the use of keys, security cards, password or pattern to open the door. The main objective is to build fully automated human facial measurements systems from images with complex backgrounds. This system can also reduce the errors which are caused by distance of object or human from the camera and by overlapping bounding boxes. The result from this research can be used in shops and mini markets to reduce losing air-condition and reduction equipment lifetime. In this paper, Viola Jones object detection framework algorithm is applied to detect the human face and tracking the detected face is accomplished by using Kanade-Lucas-Tomasi (KLT) algorithm. Moreover, the predefined bounding boxes which can eliminate the detected errors occurred in the image frame is also applied in this research. Therefore, the technique used in this system is more accurate than the others because of its precision tracker and less processing time.**

***Index Terms*— Viola Jones face detection method, Features Extraction methods, Bounding Boxes, Kanade Lucas Tomasi (KLT) method, Face Detector.**

I. INTRODUCTION

Nowadays, Automatic entrance/exit door control is widely used in public places to eliminate the need of manually opening and closing actions. Although sensors are all successful in detecting objects, they can't understand the type and the intention of the objects. For instance, a puppy or a passing pedestrian may accidentally trigger the door and cause a false opening action. Annoying, air conditioning energy waste and reduces equipment lifetime. The advantage of this system is that face tracking does not require to be touched with any hardware. Face is detected automatically by using face detection technique and the entire face tracking is completed without touching with any hardware. Face detection is the first step of the face tracking system. The performance of the entire face tracking system is influenced by the reliability of the face detection. By using face detection, it can identify only the facial part of an image regardless of the background of this image. In this system, Viola-Jones face detection method is used. Viola-Jones rescales the detector instead of the input image and run the detector many times through the image – each time with a different size. Viola-Jones has devised a scale invariant detector that requires the same number of calculations whatever the size. This detector is constructed using a so-called integral image and some simple rectangular features reminiscent of Haar wavelets job[1]. In this research, the predefined face detector is also added in order to eliminate the unrequired detected faces in the interested region. Face tracking commonly includes face detection, feature extraction and tracker initialization. This research differs from conventional door control system, human do not need to stand up in front of the camera in order to capture their faces.

The proposed system is composed of three main modules: (1) Face Detection (2) Tracking the detected human face and (3) Activating the door control system. The block diagram of proposed system is shown in Fig.1.

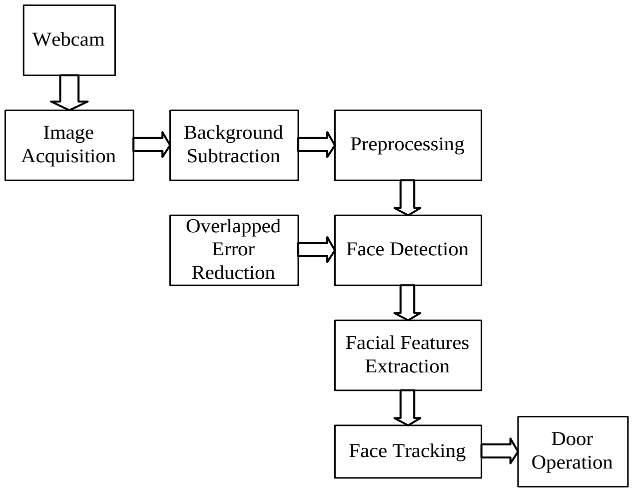


Fig.1. Block diagram of the proposed system

Automatic entrance/exit door control system by using human face is widely used in public places such as grocery stores, businesses, transportation stations, airports, and wholesale department stores. There have been various studies for human face detection and tracking in image processing. Face tracking via Spatial-Temporal Property is proposed by Jie-Ci Yang, Chin-Lun Lai, Hsin-Teng Sheu and Jiann-Jone Chen but their research based on a smart camera and also need a faster DSP module[2]. Principal Component Analysis (PCA) method is used by Shrikrishna Jogdand and Mahesh Karanjkar to implement automated door accessing system but there is a need which is to build a data base system for recognition of faces.

The first thing to keep in mind is that the Viola Jones method is a feature based detection scheme. Hamed Masnadi-Shirazi proposed his own features in order to replace the features of Viola Jones method. His method can reduce the processing time but there is not accuracy in distant

All Rights Reserved © 2016 IJSETR

image frames[3]. In this research, the more accurate bounding boxes (BBoxes) are used in order to detect only the human face in the interested region. Moreover, the size of the bounding boxes can be adjusted according to the image frame.

II. System Overview

The proposed system detects the human face, extracts the facial features from the detected face and tracks the features in each successive frame. The input of the proposed system is a human face captured by a webcam from a distance of one hundred centimetres. Firstly, pre-processing is applied on detected face image frame. And then, Viola Jones object detection method is applied on the result image for detecting only the human face after applying the pre-processing. The background subtraction is applied to know whether there is motion or not in the interested region. If there is motion in the region, detect only the facial part of the human in the image by using the Viola Jones algorithm. The facial features are extracted from the detected human face to track the motion. The door opening operation is performed after the two seconds tracking. The flowchart of the door opening system is shown in Fig.2.

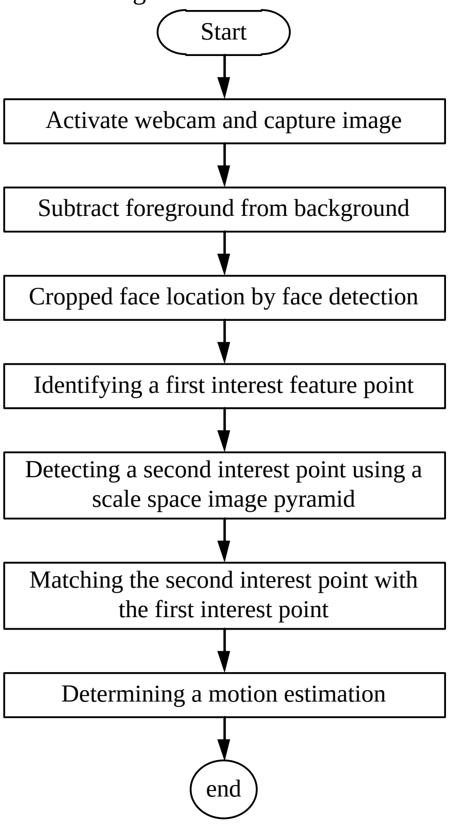


Fig.2. Program flowchart of the door opening system

III. Pre-processing

The aim of preprocessing is an improvement of image data that suppresses undesired distortions or enhances some image features important for further processing. The first step of the image processing is to subtract the foreground from the background. So, the background subtraction is firstly performed in the region of interest (ROI). After that, each inputting image is converting the gray-scaled for thresholding. And then the image is also converted into binary scaled that is black and white pixel values for image processing. If there is human in the interested region detect only the facial part of that person. The facial features are extracted from the detected face of human who is passing in front of the door. The extracted features are found in each successive frame in order to track.

A. Grayscale Image

In photography and computing, a grayscale digital image is an image in which the value of each pixel is a single sample, that is, it carries only intensity information. Images of this sort, also known as black and white, are composed exclusively of shades of gray, varying from black at the weakest intensity to white at the strongest. Grayscale images are distinct from one-bit bitonal black and white images, which in the context of computer imaging are images with only two colors, black and white (also called bilevel or binary images). Grayscale images have many shades of gray in between. The intensity of a pixel is expressed within a given range between a minimum and a maximum, inclusive. This range represented in an abstract way as a range from 0 (total absence, black) and 1(total presence, white), with any fractional values in between. Captured image (original image) and gray image are shown in Fig.3 and Fig.4.

Fig.3. Original image

Fig.4. Gray image

B. Finding object position by counting pixels

The black and white are function returns the area of a binary image. The area is a measure of the size of the foreground of the image. The area is the number of on pixels in the image. However, the black and white area does not simply count the number of pixels set to on the image. Rather, black and white area weights different pixel patterns unequally when computing the area. This weighting compensates for the distortion that is inherent in representing a continuous image with discrete pixels. For example, a diagonal line of 50 pixels is longer than a horizontal line of 50 pixels. As a result of the weighting black and white area uses, the horizontal line has area of 50, but the diagonal line has area of 62.5. In this proposed system, the foreground of the image is identified as black and background is defined by white pixels. If the is not black pixels in the image, it is noted that there is not object in the frame. Then the pixels are changed from black to white pixels if there are black pixels in the interested frame. The black and white change of the image is shown in Fig.5. The inverted binary image is show in Fig.6.Then the foreground

All Rights Reserved © 2016 IJSETR

of the image is found by counting the number of pixels. The detected object position in the frame is shown in Fig.7.

Fig.5. Binary image Fig.6. Inverted binary image

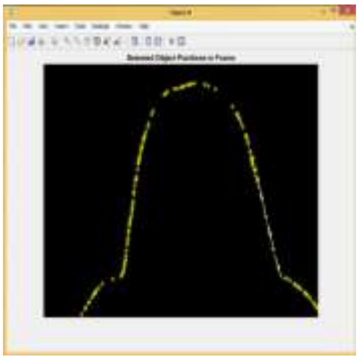


Fig.7. Detected object position in the frame

C. Merge Detection Threshold

Detection threshold is specified as a pair consisting of 'MergeDetections' and a scalar integer. This value defines the criteria needed to declare a final detection in an area where there are multiple detections around an object. Groups of collocated detections that meet the threshold are merged to produce one bounding box around the target object. Increasing this threshold may help suppress false detections by requiring that the target object be detected multiple times during the multiscale detection phase. When this property is set to 0, all detections are returned without performing thresholding or merging operation. This property is tunable. For each increment in scale, the search window traverses over the image producing multiple detections around the target object. The multiple detections are merged into one bounding box per target object. The Merge threshold property is used to control the number of detections required before combining or rejecting the detections. The size of the final bounding box is an average of the sizes of the bounding boxes for the individual detections and lies between MinSize and MaxSize. The searching window or bounding box in the image is shown in Fig.8.

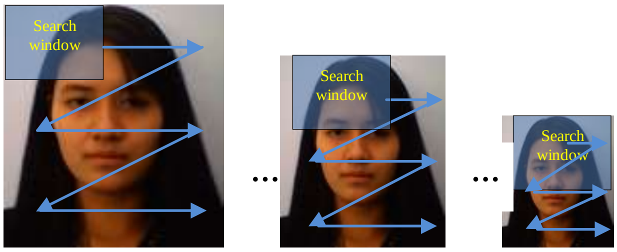


Fig.8. Searching window (or) bounding box in the image

The detector incrementally scales the input image to locate target objects. At each scale increment, a sliding window, whose size is the same as the training image size, scans the scaled image to locate objects. The scale factor property determines the amount of scaling between successive increments. The merge detection thresholds are shown in Fig.9.

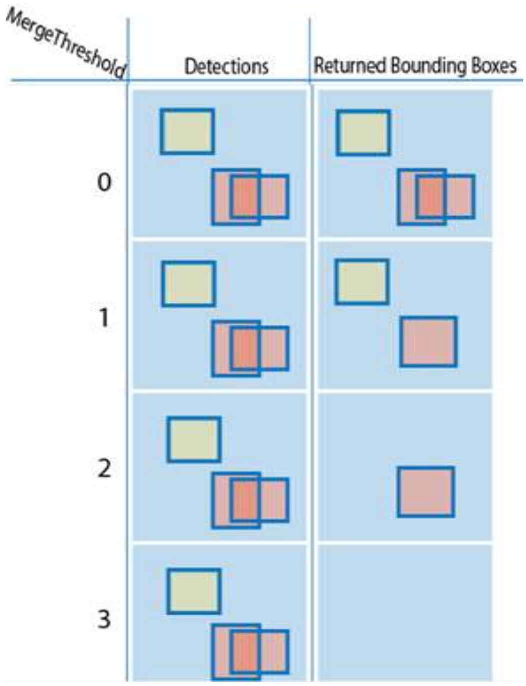


Fig.9. Merge detection threshold

Set the ratio type to 'Union' to compute the ratio as the area of intersection between bboxA and bboxB, divided by the area of the union of the two. Set the ratio type to 'Min' to compute the ratio as the area of intersection between bboxA and bboxB, divided by the minimum area of the two bounding boxes. The detected people and detect scores before merging is shown in Fig.10.

$$\text{Union} = \frac{\text{area}(A \cap B)}{\text{area}(A \cup B)} \qquad (1)$$

$$\text{Min} = \frac{\text{area}(A \cap B)}{\min(\text{area}(A), \text{area}(B))} \qquad (2)$$

All Rights Reserved © 2016 IJSETR

ISSN: 2278 – 7798
International Journal of Science, Engineering and Technology Research (IJSETR)
Volume 5, Issue 9, September 2016

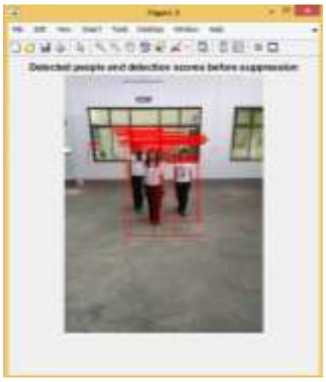


Fig.10. Detected people and detect scores before merging

The overlapped bounding boxes which are occurred in detection system can be eliminated by using merge threshold. To remove the unused bounding box the classification threshold are changed. By using this changing process, the errors which are occurred in detection system can be removed. The detected scores and detected people after suppression is shown in Fig.11.

Fig.11. Detected people and detect scores after merging

The size of the face detector can be changed in order to reduce the errors in image. And also the threshold value of detector is changed to eliminate the detection errors. The results of removing unrequired detections can be seen in Fig.12 and Fig.13.

Fig12. Elimination errors in detected image

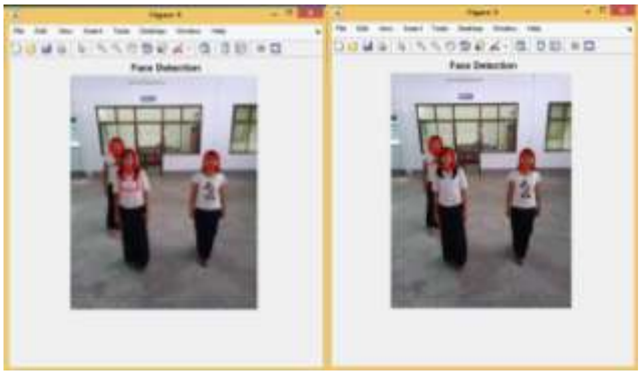


Fig.13. Elimination errors in detected image

In Fig.12, the minimum size of the detector is [50 50] and the merge threshold value is 10. By using these values, the overlapping bboxes which are occurred in face detection system of the image. And also the face detector minimum size [20 20] and face detector merge threshold 9 is applied to eliminate the detected errors in Fig.13. The more the value of mere detection threshold the less the overlapped bounding boxes errors. After the elimination of error in the detected frame, the successful frames are tracked about two seconds in the interested region in order to activate the door operations.

IV. Tracking System

Kanade–Lucas–Tomasi (KLT) feature tracker is an approach to feature extraction. KLT makes use of spatial intensity information to direct the search for the position that yields the best match. It is faster than traditional techniques for examining far fewer potential matches between the images. This approach keeps track of the face even when the person tilts his or her head or moves toward or away from the camera [4]. In a simple KLT algorithm generally consists of the following basic steps:

1. Detect Harris corners in the first frame.
2. For each Harris corner compute motion between consecutive frames.
3. Link motion vectors in successive frames to get a track for each Harris point.
4. Introduce new Harris points by applying Harris detector at every frame.
5. Track new and old Harris points using steps 1-3.

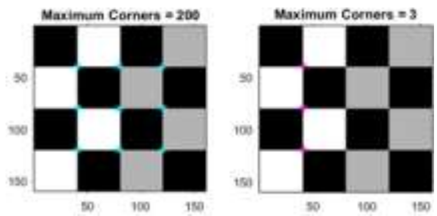
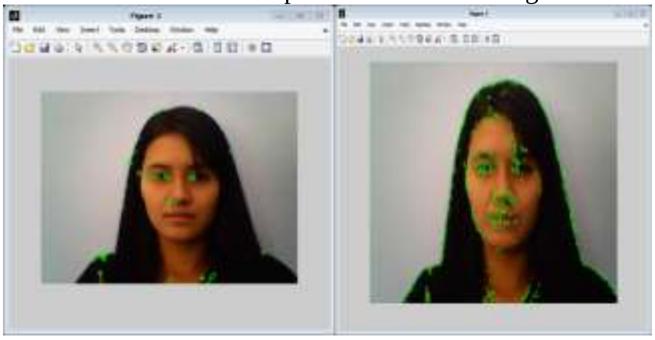


Fig.14. Different numbers of Harris corners

The feature extraction with different maximum corner points in a chessboard is shown in Fig.14. The corners are useful image features because they are necessarily distinct from their neighboring pixels. The corner is the intersection of two edges. The Harris corner detector is a standard algorithm for corner detection in computer vision. The algorithm works by analyzing the eigenvalues of the 2D discrete structure tensor matrix at each image pixel and flagging a pixel as a corner when the eigenvalues of its

All Rights Reserved © 2016 IJSETR

structure tensor are sufficient large. Intuitively, the eigenvalues of the structure tensor matrix associated with a given pixel describe the gradient strength in a neighborhood of that pixel. As such, a structure tensor matrix with large eigenvalues corresponds to an image neighborhood with large gradients in orthogonal directions that is a corner. The results of different corner points are shown in Fig.15.



(a)Strongest Corners 50 Point(b)Strongest Corners 500Points

Fig.15. Different numbers of corner points in the same image

The aim of the KLT algorithm is to find the coordinates in the image which has a varying texture. First of all features have to be selected which can be tracked from image to image in a video image stream. The selection of the features is based on texture (intensity information). A number of fixed-sized feature windows are selected on the first image of a sequence. These feature windows are tracked from one image to the next using the KLT method. This method calculates the sum of squared intensity differences between a feature in the previous image and the features in the current image. The displacement of the specific feature is then defined as the displacement that minimizes the sum of differences.

In KLT tracking method consists of three steps: object detection, tracking objects from frame to frame and trajectory estimation. In this system, multiple object tracking is used to find the trajectory of the target objects through a number of frames from an image sequence. The object tracking is not only important but also a very difficult because of the arbitrary object shapes, illumination changes, object occlusion, complex object shape and motion, and camera motion[5].

V. Simulation Tests And Results

In this paper, there are two parts for implementation steps. The first is the implementation of face detection by using Matlab and the second is the implementation of Kanade-Lucas-Tomasi algorithms for tracking. Firstly the image is captured from the live stream through the web cam. By making background subtraction the captured image from the original background image, the motion in the sense can be detected. If there is motion detection in the interested region, then perform the preprocessing step with personal computer (PC). Moreover, the merge detection thresholding function is also added in this system in order to eliminate the unrequired detected faces. The detection errors are caused by the overlapped bounding boxes or search windows in the interested region. In this system viola jones algorithm is mainly affected to detect the human face who is passing in front of the door. This viola jones algorithm is applied to the detected object in order to find the position of the face in the interested region. If the detected object is human, then Kanade-Lucas-Tomasi algorithms is applied to that object to track the motion of face. By using KLT algorithm in this research, the facial features can be tracked successfully in different lighting condition. The results in illumination changes can be seen Fig.16 and Fig.17.

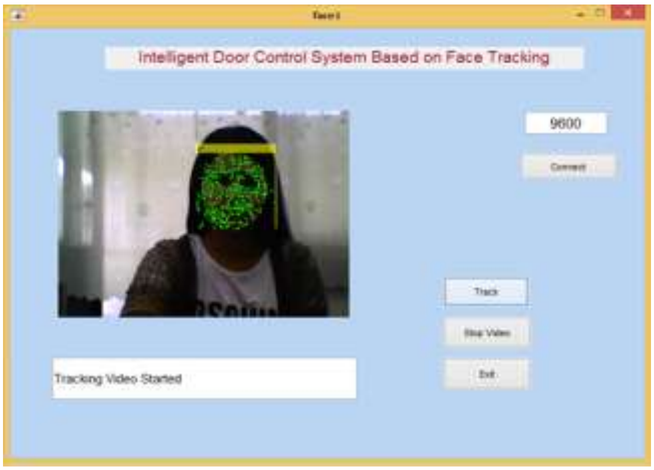


Fig.16. Face tracking in the dark condition

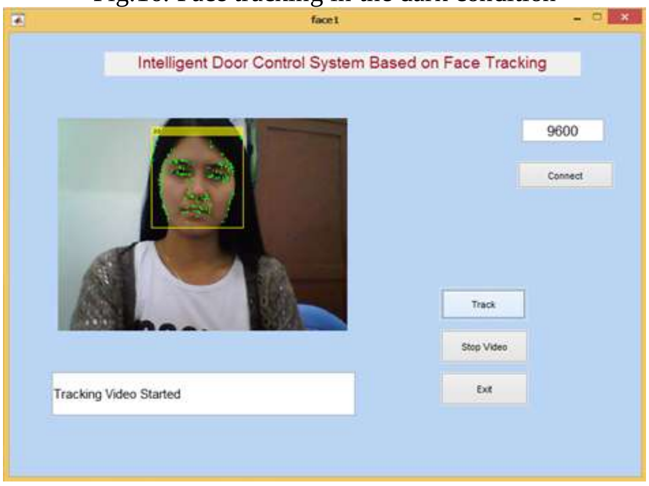


Fig.17. Face tracking in the lighting condition

This is a system for detecting and tracking multiple faces from live video through the webcam in order to accomplish the intelligent door control system. The system detects faces using the Viola Jones algorithm, detects min-eigen corners within each face's bounding box, and tracks the corners using Kanade-Lucas-Tomasi (KLT) algorithm. It re-detects the faces every 2 frames in order to correct the tracker and replenish the points. The multiple tracking of the faces in the image sequence are shown in Fig.18 and Fig.19.

All Rights Reserved © 2016 IJSETR

ISSN: 2278 – 7798
International Journal of Science, Engineering and Technology Research (IJSETR)
Volume 5, Issue 9, September 2016

Fig.18. Matlab GUI result when the two persons are detected

Fig.19. Multi tracking in the live stream

In Fig.18, the two faces are tracked in the interested region and the door can activate opening operation for that two persons. Moreover, the more people can also be tracked in the region of interest (ROI) and that result is illustrated in Fig.19. For this implementation, facial features must be extracted from the detected face firstly. For each feature point in the previous frame, the point tracker will now attempts to find the corresponding point in the current frame. The door opening system is started after the two seconds tracking of face who has intention to enter to the door. When people have the intention to pass through the door, they must look at to the door. At that time, the detecting and tracking of face can be done easily. After that, the door operation system can be performed. The door opening action of the system is simulated in Fig.20.

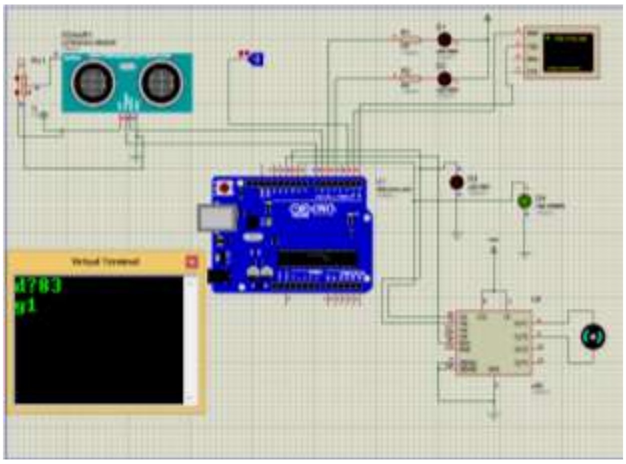


Fig.20. Simulated result for door opening action

In this system, the simulation is connected with Matlab program in personal computer (PC). Arduino and Matlab both know the rate (BAUD RATE) at which they share infos and the physical serial port they are talking through, otherwise they are not going to understand each other. In this system, Matlab and Arduino microcontroller is connected with 9600 baud rate in each. The range of the interested region to find the human face is limited into hundred centimeter form the camera. After the human who has intention to pass through the door is detected by using the ultrasonic sensor in the limited range, the detected data is sent to the Matlab program in order to process face detection and tracking implementations. If the detection and tracking operations are successful, the commands are sent to the Arduino microcontroller from the Matlab program. The door opening operation is activated by turning the door motor forward if the data 'g1' is received. If the data 'g0' is sent from Matlab to Arduino, the door closing action is activated by turning the door motor in reverse direction. After activating the door opening action, the sensor is used to prevent people from being trapped by the door before they leave the passage. If there is not person who has the intention to pass through the door in front of the door, the door will close automatically.

VI. Conclusion

In this paper, automatic door access system by using face detection and tracking is presented. The automatic face detection and tracking is done by Matlab program on PC. Door is opened immediately after confirming that the person has attention to pass through the door. To detect the face in the image, I have used a face detector based on the Haar-like features. This face detector is fast and robust to any illumination condition. Viola-Jones face detection method is used to detect the location of the face in an image. Since this detection method can detect only face images for frontal view correctly, this system has limitations in head orientation. Kanade-Lucas-Tomasi method is used to extract the important facial features and to track the face within two seconds. Using detected points with the algorithm of Shi and Tomasi, good results in video sequence and in real time acquisition were obtained. Therefore, this system can be used in automatic door opening in shops, stores. The proposed system has the advantages of high precision, safety, reliability, and can be responsive to demands, while

All Rights Reserved © 2016 IJSETR

preserving the benefits of being low cost and high added value.

ACKNOWLEDGMENT

The author would like to express her gratitude to all her teachers from Department of Mechatronics Engineering, Mandalay Technological University. The author wishes to express special thanks and sincere gratitude to her supervisor, Dr. Wut Yi Win, Associate Professor and Head, Department of Mechatronic Engineering, Mandalay Technological University for valuable suggestion, generous guidance, instructions and supervision throughout the research work. And, she deeply thanks to her parents for their support and encouragement to attain her destination without any trouble. The author also especially thanks to her friends for their supports and kindness.

REFERENCES

[1] Paul Viola,viola@merl.com and Michael Jones, mjones@crl.dec.com, Rapid Object Detection using a Boosted Cascade of Simple Features, Accepted Conference On Comuter Vision And Patteren Recognition 2001.

[2] Jie-Cie Yang,Chin-Lun Lai, Hsin-Teng Sheu and Jiann-Jone Chen, An Intelligent Automated Door Control System Based on Smart Camera, ISSN 1424-8220, www.mdpi.com/journal/sensors, Sensors2013,13,5923-5936;doi:10.3390/s130505923.

[3]Hamed Masnadi-Shirazi, Department of Electrical and Computer Engineering University of California, San Diego La Jolla, California, hmasnadi@ucsd.edu, AdaBoost Face Detection.

[4] Kapil,B.Tech Scholar and Anurg Singh, Assistant Professor, Department of ECE, Om Institute of Technology and Management, Guru Jambeshwar University of Science and Technology, Hisar, Haryana, India.

[5] Paul Viola, viola@microsoft Way, Recomond,WA 98052,USA and Michel J.Jones, mjones@merl.com, Mitsubishi Electric Research Laboratroy,201 Boardway, Cambridge,MA 02139,USA, Robust Real-Time Face Detection, Received September 10,2001; Revised July 10 2003;Accepted July 11 ,2003.

All Rights Reserved © 2016 IJSETR