7th IEEE International Conference Humanoid, Nanotechnology, Information Technology
Communication and Control, Environment and Management (HNICEM)
The Institute of Electrical and Electronics Engineers Inc. (IEEE) – Philippine Section
12-16 November 2014 Hotel Centro, Puerto Princesa, Palawan, Philippines

Low Cost Smart Security Camera with Night Vision Capability Using Raspberry Pi and OpenCV

Wilson Feipeng Abaya, Jimmy Basa, Michael Sy
Electronics and Communications Engineering Department
Gokongwei College of Engineering
De La Salle University
Manila, Philippines

Alexander C. Abad and Elmer P. Dadios
Electronics and Communications Engineering Department
Gokongwei College of Engineering
De La Salle University
Manila, Philippines
alexander.abad@dlsu.edu.ph, elmer.dadios@dlsu.edu.ph

***Abstract*— In order to further maintain peace and provide security to people now a day, Closed-circuit television (CCTV) surveillance system is being utilized. This study focused on the design and implementation of a low cost smart security camera with night vision capability using Raspberry Pi (RPI) and OpenCV. The system was designed to be used inside a warehouse facility. It has human detection and smoke detection capability that can provide precaution to potential crimes and potential fire. The credit card size Raspberry Pi (RPI) with Open Source Computer Vision (OpenCV) software handles the image processing, control algorithms for the alarms and sends captured pictures to user's email via Wi-Fi. As part of its alarm system, it will play the recorded sounds: "intruder" or "smoke detected" when there is a detection. The system uses ordinary webcam but its IR filter was removed in order to have night vision capability.**

***Index Terms*— Background Subtraction, Haar Cascade, Human Detection, IR Modified Webcam, Raspberry Pi, OpenCV, Smoke Detection.**

I. Introduction

Closed-circuit television monitoring system has now become an indispensable device in today's society. Supermarkets, factories, hospitals, hotels, schools, and companies are having their own CCTV system for 24/7 monitoring. It gives real-time monitoring, provides surveillance footage, and allows the authorities have evidences against illegal activities. It is believe that CCTV can deter crimes. Although surveillance camera records video and helps the authorities to identify the cause of an incident such as crime or accident, it is just a passive monitoring device. The researchers developed an active surveillance camera that has the capability of identifying the context of the scene being monitored and able to give notification or alarm as the event happens.

When dealing with the real-time imaging processing, Open Source Computer Vision (OpenCV) software, a powerful library of image processing tools, is a good choice. The library is written in C and C++ and runs under Linux, Windows and Mac OS X. OpenCV is a free software that can help optimize code for basic image processing infrastructure [1]. Through the use of this software, surveillance system will not just merely recording videos, but also able to identify or differentiate the context of the scene being monitored. A good smart active low cost camera must also be capable during night time. An ordinary webcam with its infra-red (IR) filter removed can be utilized for night vision sensing with the aid of IR Light Emitting Diode (LED) illuminator.

II. Theoretical Framework & Design Consideration

A. Top Level Design

Figure 1 shows the block diagram of the whole system. The heart of the system is a single board credit card size computer known as Raspberry Pi (RPI). It is the central platform for image processing and signal alerting. An ordinary webcam with its IR filter removed captures the scene to be processed by the RPI. A dark activate switch triggers IR LED illuminator circuit during night time to enhance night vision mode. The system has the capability of detecting motion of an object by the background subtraction algorithm. When the moving object is detected, the system can classify it as human or smoke. If there is detection, the system sends out alerting signal in the form of sound and email to indicate fire or unauthorized invasion.

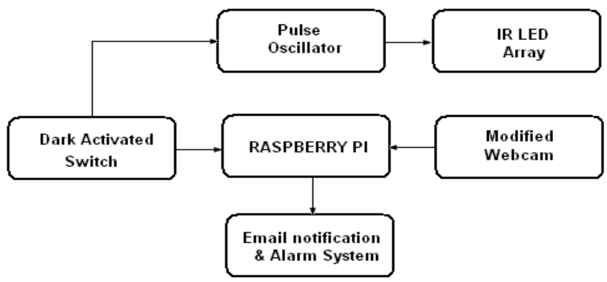

Fig. 1. System Block Diagram

B. General System Flow Chart

To optimize the algorithm, the researchers separate the algorithm into three parts which are motion detection, human detection and smoke detection. If there is no motion detected, the program will not go to human detection and smoke detection algorithm. Otherwise, if motion is detected, the current frame of detected motion will be processed by human detection algorithm. If the moving object is identified as non-human object, the smoke detection algorithm will be activated. If there is a detection of either human or smoke, alarm, LED indicators and email interface function will be activated. General system flowchart is shown in Figure 2.

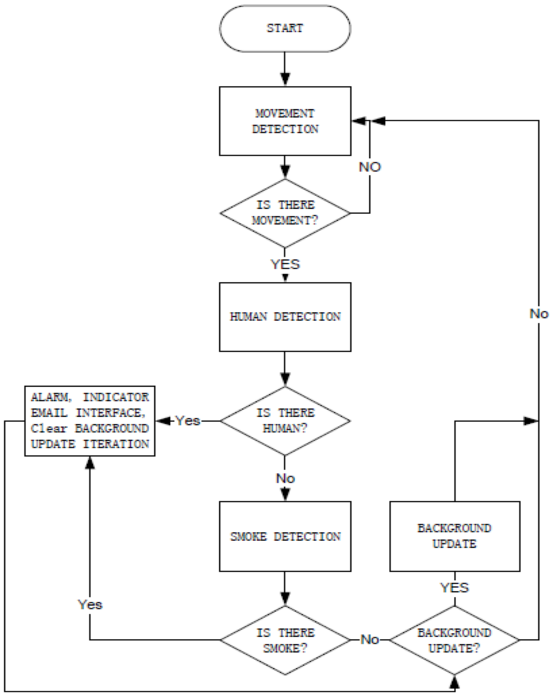

Fig. 2. General System Flowchart

C. Background Subtraction for Motion Detection

Since the camera used is monitoring at a fixed location, background subtraction algorithm can be used in detecting motion by the concept of frame differencing. Moving objects within a given background can be deduced from the difference of the current frame and a reference frame, often called the "background image" [2].

There are several algorithms that can be used for motion detection like optical flow and edge analysis. One of the simplest forms of motion detection is using background subtraction. There are different types of background subtraction based from the different journals. The challenges in most of these algorithms are the illumination changes and processing speed. In choosing the type of background subtraction that will be used in the system, the researchers considered the place where the system will be placed and also the processing speed of the RPI. According to one comparative study, the basic background subtraction has the lowest processing speed requirement but it can't be implemented in a complex background environment [3]. The researchers finally choose the basic background subtraction for motion detection due to its low computational requirement. Figure 3 is the block diagram of the simple background subtraction algorithm which can be implemented using equation (1).

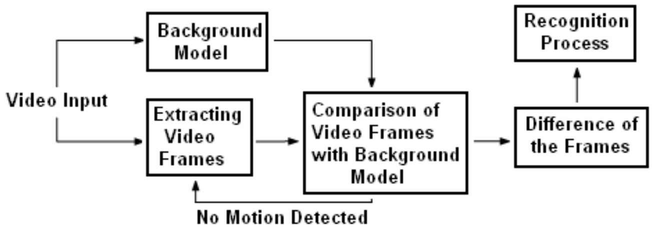

Fig. 3. Block Diagram of Background Subtraction

$$|Image(x,y,t) - Image(x,y,t-1)| > Th \quad (1)$$

In the equation 1, each pixel that occurs in the coordinate(x,y) of the current image(t) is subtracted by the previous image(t-1). If their difference reached a threshold value (Th), the motion is detected.

D. Human Detection

Human detection part fully depends on the program of background subtraction. If there is no motion detected, LED indicator will not light up; then the program will start to update the background. On the other hand, if motion detected, that particular frame will be the input frame to the human detection process and indicator for motion detect will light up. The program will immediately convert the captured image to gray scale, improve the contrast of the image and store it in the memory (this is necessary in order to improve the processing speed of the program). The converted image will be process through Haar Classifier Stages. If head and shoulder feature is not detected, the program considers it as non-human object and will move to smoke detection; otherwise, the program will send the captured image to the user's email and triggers the sound alarm and LED indicator.

Haar like features [4] is one of the common ways used for object detection. Just like the human detection method presented in [4], the researchers used head and shoulder feature as the shape or "region of interest" (ROI) rather than using the typical face detection. It is because basic shape like face in an image can occur in varying scale, position and orientation. In detecting human, the task becomes more challenging owing to the largely varying size, shape, posture and clothing. That is why the researchers considered head and shoulder to be the more appropriate since it is the most unvarying part of human body [5]. Fortunately, OpenCV has available haar cascade file "haarcascade_mcs_upperbody.xml" covering head and shoulder which is used in this study. The range of monitoring in this study is up to five meters. If face detection is used, it is difficult to detect face at a long distance since the face will be small and blurry. On the other hand, head and shoulder is still quite viewable when the object is 5 meters away from the camera. Table 1 shows the different types of motion sensing classifiers [6].

TABLE I. DIFFERENT TYPES OF CLASSIFIERS [6]

Paper	Human Model	Classifier
Cutler and Davis [2000]	Periodic Motion	Motion Similarity
Utsumi and Tetsutani [2002]	Geom. Pix. Val.	Distance
Gavrila and Giebel [2002]	Shape template	Chamfer Dist.
Viola et al. [2003]	Shape + Motion	Adaboost Cascade
Sidenbladh [2004]	Optical Flow	SVM (RBF)
Dalal and Triggs [2005]	HOG	SVM (Linear)

Cutler and Davis show a technique on detecting periodic motions. Utsumi and Tetsutani show a method that is based on the fact that the relative positions of various body parts are common to all humans, despite the unstable factors like illumination. Gavrila and Giebel's paper focus on the scenario of a moving camera mounted on a vehicle. Viola's method is probably the most familiar classifier, it is able to detect human base on shape and motion features. Sidenbladh's technique is made by collecting several samples of human and non-human motion and computing optical flow. Dalal and Triggs's method uses the fact that the shape of an object can be well represented by a distribution of local intensity gradients or edge directions. [6]

Paul Viola and Michael Jones's [4] research papers have a significant improvement on image processing. On their paper entitled "Rapid object detection using boosted cascade of simple features" in 2001, they introduced a method that can accurately detect object using Haar feature-based cascade classifiers. These features use the change in contrast values between adjacent rectangular groups of pixels. The relative light and dark areas are being determined by the contrast variances between the pixel groups. Two or three adjacent groups with a relative contrast variance form a Haar-like feature. Equation 2 and 3 are used for computing integral areas needed for developing Haar Classifier [7].

$$AI[x,y] = \sum_{x'\le x, y'\le y} A[x',y'] \tag{2}$$

Where A[x', y'] = the original integral image
AI[x, y] = integral image

$$AR[x,y] = \sum_{x'\le x, x'\le x-|y-y'|} A[x',y'] \tag{3}$$

Where A[x', y'] = the original integral image
AR[x, y] = rotated integral image

E.Smoke Detection

In order to develop the algorithm that will be used for the smoke detection, the researchers considered the behavior of the smoke. One of the main behaviors of the smoke is that when smoke begins to scatter, the area of the moving area will gradually increase. The direction of hot smoke is always upward. Since the system is being utilized indoor, the external factors like wind will not affect the movement of the smoke. In getting the direction of the movement of the smoke, the researchers make use of the moment extracted from the motion detection algorithm to obtain the coordinate of the centroid. Refer to (4) and (5) for getting the moment and the Y-centroid. If the y-coordinate of the centroid is decreasing, it means that the area of the contour drawn is going up.

In the first part of the smoke detection algorithm, researchers sum up the areas obtained from the motion detection algorithm and get the average Y-centroid value. If the relationship of the area and Y-centroid of the current frame and the previous frame satisfy the given condition, there will be an increase in y value. The y value is used to indicate the iteration of smoke detection. When y value reaches a certain threshold, it means there is a smoke. When there is a smoke, the alarm, indicator and email feature will turn on. If there is no smoke and Y-centroid value reaches two, the iteration for the background update will reset to zero.

$$m_{ji} = \sum_{xy} (array(x,y) \cdot x^j \cdot y^i) \tag{4}$$

$$\bar{y} = \frac{m_{01}}{m_{00}} \tag{5}$$

Where: m_{ji} is the moment
$\bar{y}$ is the center of mass in y-axis

F.Raspberry Pi

RPI Model B was utilized in the researcher's project and it has 512MHz SDRAM, and 700MHz CPU which allows a typical computing environment. The external ports of RPI being used in this research are the GPIO port, audio port and Ethernet port. GPIO and audio ports are used for alarming system and Ethernet port was used to connect RPI to the monitoring device or to the networking device. Shown in Figure 4 is the picture of RPI.

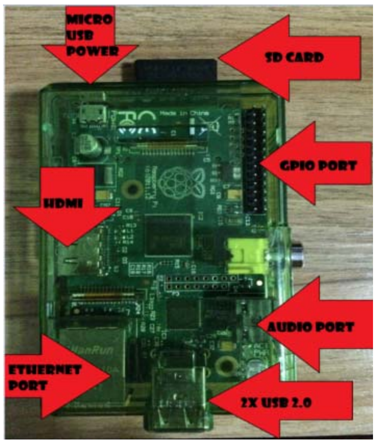

Fig. 4. The External Ports of Raspberry Pi

G. Dark Activated Switch

For the light sensing circuit, the researchers use LDR and an inverting voltage comparator circuit to act as a switch to the oscillating circuit and input of RPI. A reference voltage was put on the non-inverting input of the op-amp using potentiometer. The sensitivity of the circuit can be adjusted using potentiometer. The voltage divider circuit compose of a LDR and a resistor is used as the input signal of op-amp. When the LDR is exposed to light, the voltage across the R1 will increase since the resistance of the LDR is very low. This means the input voltage is greater than the reference voltage; therefore the output will be low. On the other hand, the resistance of the LDR increases if there is no presence of light. This cause voltage drop across R1 and the output of the Op-amp will be high. The output of the Op-Amp is used as the trigger to the transistor and RPI.

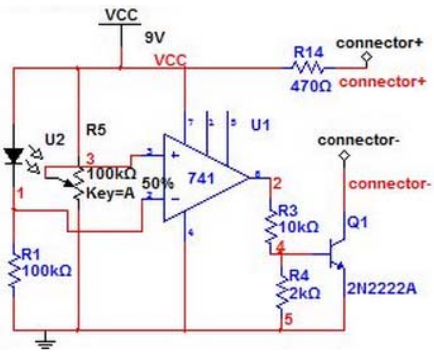

Fig. 5. Dark Activated Switch

H. Oscillating Circuit and IR illuminator

The researchers used an oscillating circuit to turn on and off the Infrared LEDs. It is because in order to make the LEDs to emitting out the max light it can handle, max current is required. In normal operating mode, the maximum continuous current for IR LED is 100mA, but the peak pulse current is about 1A, that is why an oscillating circuit is need, to supply the current as much as possible (not exceeding 1A) while not destroying the LEDs. The target current flowing through the LEDs is about 400-500mA and operating it at a frequency of 10 kHz. The desired duty cycle is 10% to protect the LEDs from bursting. The circuit also reduces power consumption.

Since the 555 timer cannot produce a duty cycle less than 50%, the researchers designed it to be 90%, and then used an inverter to achieve the desired 10% duty cycle.

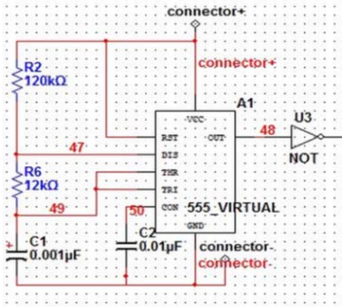

Fig. 6. Oscillator

I. Sound Alarm and Email Feature

In this research paper, the researcher added an email notification feature since RPI itself already had Simple Mail Transfer Protocol (SMTP) server embedded. This feature will be activated once the human or smoke had been detected. The RPI will sent an email with the corresponding attachments to the receiver.

The RPI also had a feature that could play audio file so the researcher had utilized this function such that the RPI could output the corresponding audio file through its audio port based on what had been detected.

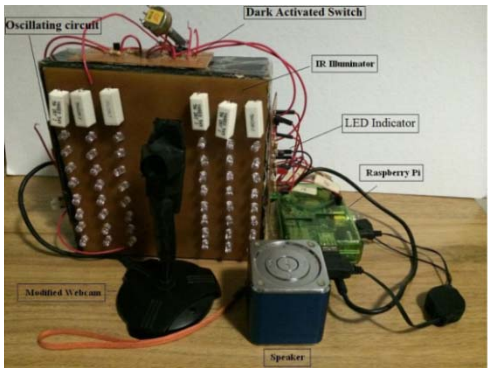

Fig. 7. Overall Set-up of the System

Figure 7 showed the integrated system of the thesis which compose of IR illuminator which served as the light source at night time, raspberry pi which served as the central processing unit, modified webcam which served as the sensor for human and smoke detection, oscillating circuit which served as switch for IR illuminator, dark activated switch which served as an automatic switch to shift from day to night vision and lastly LED indicator and speaker which served as the alarm system.

III. Results and Discussion

Night Vision Output

Figure 8 shows the difference and the actual output of the IR array.

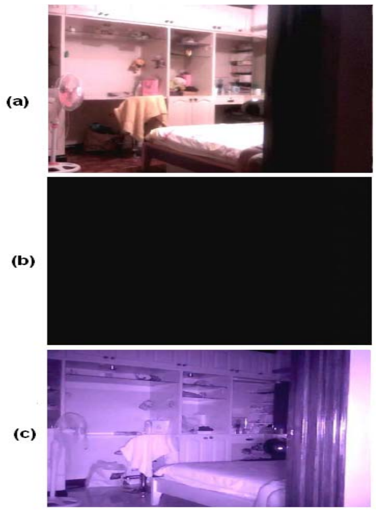

Fig. 8. (a) Room under normal light condition. (b) the light is switched off. (c) the actual lighting condition when IR array is turned on

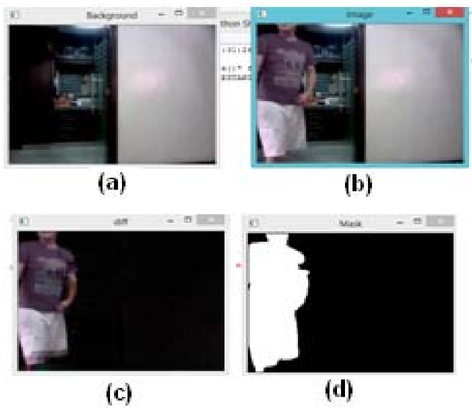

Fig. 9. Background Subtraction

Figure 9 shows the output of the background subtraction. Figure 9a is the background. The moving object was successfully being identified and isolated by the algorithm as shown in Figure 9b-c.

Figure 10 shows the positive samples are being detected by the system, even if the person is covered by the mask, wearing a hat or covered by a piece of cloth, as long as the head and shoulder feature remains, the system will be able to detect it. Table 2 shows the test results under day time and night time. The total accuracy for human detection is 83.56%.

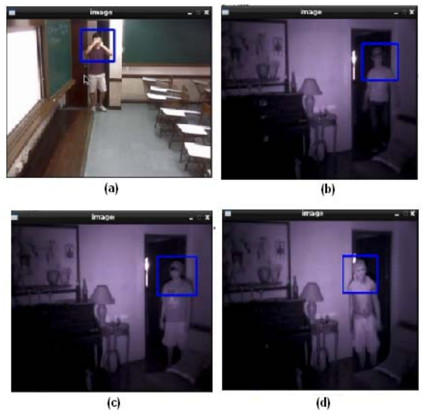

Fig. 10. Human Detection Samples at Day Time and Night Time

TABLE II. HUMAND DETECTION RESULTS FOR 27 VIDEOS

Trials	True Positive	False Negative	False Positive	True Negative	ACCURACY
1-6(Day)	8	0	1	2	90.90%
7-8(Day)	7	0	2	1	80%
9(Day)	6	0	1	0	85.71%
10(Day)	6	0	2	0	75%
11(Day)	3	0	2	3	75%
12(Day)	11	0	1	1	92.31%
13(Day)	5	0	0	0	100%
14(Day)	1	0	0	0	100%
15(Day)	11	0	0	1	100%
16(Day)	5	1	1	1	75%
17(Day)	6	0	1	1	87.5%
1-4 (Night)	10	0	2	0	83.33%
5(Night)	4	0	1	1	66.67%
6(Night)	5	0	1	0	83.33%
7(Night)	4	1	3	0	50%
8(Night)	5	0	1	0	83.33%
9(Night)	5	0	1	0	83.33%
10(Night)	4	1	0	0	80%
11(Night)	6	0	0	0	100%
12(Night)	6	0	1	1	87.5%

Figure 11 show the smoke detection during day time. The behaviors of the smoke were being successfully point out as showed in the blob of the Figure 11.

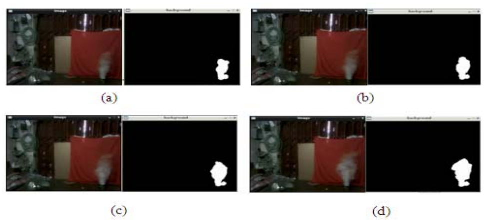

Fig. 11. Smoke Detection Day Time

Figure 12 show the smoke detection during night time. The system was able to successfully capture the movement of the smoke although the video was in black and white color base.

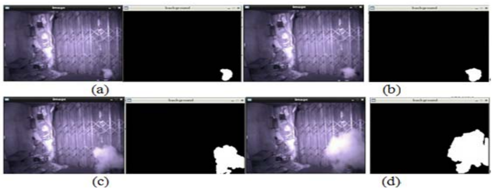

Fig. 12. Smoke Detection Night Time

The moving pixels in the sample videos above were successfully being masked out. The contours of the sample video showed that there was an actual increased in the area and an uprising movement was also being observed by the researcher. Table 3 shows the test result of the smoke detection under day and night time. The detecting results of the 12 sample videos were able to have 83.33% accuracy.

TABLE III. SMOKE DETECTION RESULT FOR 12 TEST VIDEOS

Trials	True Positive	False Negative	False Positive	True Negative
1(Day)	1	0	0	0
2(Day)	1	0	0	0
3(Day)	1	0	0	0
4(Day)	0	1	0	0
5(Day)	1	0	0	0
6(Day)	1	0	0	0
7(Night)	1	0	0	0
8(Night)	0	1	0	0
9(Night)	1	0	0	0
10(Night)	1	0	0	0
11(Night)	1	0	0	0
12(Night)	1	0	0	0

To test the response time of the RPI, the processing speed in frame per second (fps) of the RPI was being measured. The RPI could only processes up to 0.4 fps and this is already being processed by RPI with the maximum over clocked speed of 1GHz. Although there was few seconds of delay for the detection, it is said to be real time because the RPI was able to provide necessary response few seconds after the event.

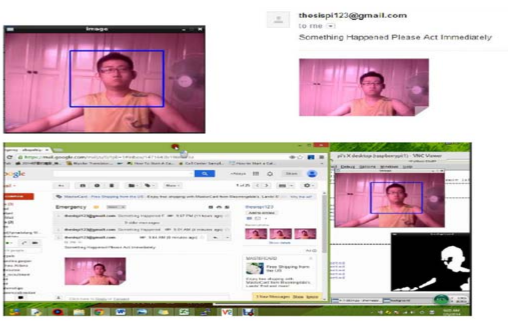

Fig. 13. Email Interface

IV. CONCLUSION AND RECOMMENDATION

In this paper, the researchers were able to create a ₱7,000 (PHP) smart security camera that has human detection and smoke detection capability. When compared to a smart camera set-up with specialized IR camera and laptop as its processor, the total cost of this project is much cheaper. Night vision capability is achieved by using IR LED illuminator and a modified webcam. The accuracy for human and smoke detection is 83.56% and 83.33% respectively. For the alarm system, RPI was utilized to output the designated audio sound to the speaker and sent the captured positive detection images via email which notified the user and provide the user essential time to prepare the upcoming event. For the researchers, safety is the primary concern and they believe an integrated smart camera system will be a significant help in the security.

For those who would like to develop the system presented here, there are other single board computers aside from Raspberry Pi that can be used as a substitute for faster processing but at a higher price.

ACKNOWLEDGMENT

The authors would like to thank De La Salle University for helping us to finish this project.

REFERENCES

[1] Ali A.; Farzin F. (2001). “An Implementation on Object MoveDetection Using OpenCV,” Department of Systems and computer Engineering, Carleton Univ., Ottawa.

[2] M. Piccardi (October 2004). "Background subtraction techniques: a review," IEEE International Conference on Systems, Man and Cybernetics 4: 3099–3104

[3] Y. Benezeth; P.-M. Jodoin; B. Emile; H. Laurent; C. Rosenberger4. (2010)," Comparative Study of Background Subtraction Algorithms". Journal of Electronics Imaging19 (2010). Oct. 16, 2012

[4] Viola, P. and Jones, M. (2001). “Rapid object detection usingboosted cascade of simple features,” IEEE Conference on Computer Vision and Pattern Recognition, 2001.

[5] Mukherjee, S and Das, K. (2013). “A Novel Equation basedClassifier for Detecting Human in Images,” International Journal of Computer Applications (0975 – 8887), Volume 72– No.6, May 2013

[6] Ogale, N. (2006). “A survey of techniques for human detectionfrom video,” Department of Computer Science Maryland Univ. 2006.

[7] Wilson, P. and Fernandez, J. (2006). “Facial feature detectionusing Haar classifier”. Journal of Computing Science in Colleges, Volume 21 Issue 4, Pages 127-133, April 2006.