

Real Time Vision Based Security System

Sachin Patidar^{#1}, Amit Prakash Pandey^{*2}, K. Ketan^{\$3}, Gajjar Rushi
Pareshkumar^{&4}

[#] M.Tech. II Year (School of Electronics Engineering, VIT University, Vellore, Tamilnadu – 632014 INDIA)

^{*} M.Tech. II Year (School of Electronics Engineering, VIT University, Vellore, Tamilnadu – 632014 INDIA)

^{\$} M.Tech. II Year (School of Electronics Engineering, VIT University, Vellore, Tamilnadu – 632014 INDIA)

[&] M.Tech. II Year (School of Electronics Engineering, VIT University, Vellore, Tamilnadu – 632014 INDIA)

Abstract: Motivation for the Real time vision based security system arises due to increasing crime activity now a days, which includes intrusion in high security areas where the presence of an unauthorized person is unacceptable. The paper carries out research on the low cost implementation of a system for change detection in a closed monitored area. This paper presents a vision based security system that includes a Linux based target board, web-camera and indicators to alert the concerned user. This work is successfully developed on embedded target platform board comprising of a Broadcom multimedia processor BCM2835 based on ARM11 embedded processor core as the hardware and the software is implemented on the Linux kernel by porting cross-compiled OpenCV, GPU and GUI libraries as well as UVC drivers for USB camera. An image is captured from a web-camera at random intervals and using change-detection technique i.e. image subtraction, thresholding and morphological operations the final image with major blob is produced. The area of the major blob is derived by calculating mean of pixel array elements. These values of cross section area are compared with calculated values. If derived values fall in the range of theoretical values, it can be identified as human intrusion. These ranges of theoretical values (the average cross sectional area of human) are defined in Anthropology theory. This system has been tested for different environments, such as an experimental closed monitored area, and achieves over 86% detection rate at 5-6 frames/s processing speed.

Keywords: ARM11, Motion Detection, Morphology, OpenCV, Random Image Update.

I. Introduction

This is an era of technology and intelligent electronics system designs. The use of vision based and wireless devices make home and work-area smarter. It is necessary to build a real time security system which can make living more convenient, benign as well as comfortable. For home and office safety, automated security systems play a significant role of safeguarding the area (of interest). These systems also ensure that intrusions are disillusioned and intruders are sensed as well as tracked in a closed supervised area. People need such systems to track an unintentional event when they are absent at their place. This system ensures the safety of the place or a high security area.

A lot of research work has been done on different type of security systems to make the area, intrusion free. John et al. [1] proposed a novel algorithm which combined the face recognition and motion detection module using primary and secondary processor system (Multi-processor system) for a purpose of detecting an intruder. In 2014, Donabel et al. [2] introduced remote farm security systems which ensured the security of crops in a field; the system used the OpenCV library for implementation with motion detection and image subtraction algorithm. In 2005, Luo et al. [3] introduces a moving robot in a closed supervised area with the gas and fire sensors deployed, this system had been implemented to safeguard the homes from fire and gas leakage. Had a camera module been implemented on this unit, the surveillance can be done from remote place. Qusay [4] implemented the real time home security system implemented on Java technology which detected and marked the change on the refreshed images and it was able to send an SMS to the mobile phone user and it also sent an email to a user if the system is connected with the internet.

The proposed system works on this methodology: Initially when no one is present in the closed supervised area, the frame is segmented from the video captured by the camera, which we call as a reference image or background image. At a random interval we segment the video to single frame and capture the current image. The currently captured image is compared with the reference image using methodologies used in section IV. The system also updates the reference image at random intervals which makes the system adaptive to illumination changes which happen throughout the day. If the changes are detected, the proposed system can alert the concerned user via different methods. This system is implemented using a single module with a web camera connector and a memory card, which takes the backup of the refreshed images on a memory chip interfaced with the system ensuring the safety of data in absence of power.

In future implementation the system can be upgraded to alert user via internet or mobile services.

The vision based security system has numerous advantages. First and foremost, it is easy to install. It can be easily setup through simple wiring and can be interfaced with most of the processing nodes. It is easy to interpret and user friendly. The benefit of this implemented system is that, it can be easily integrated with an existing security system. The OpenCV implementation in C language makes it compatible to upgrade almost all existing systems. The existing systems just need to be ported with OpenCV libraries and the implementation firmware. Being easy to set up, it is inexpensive and unobtrusive. It eradicates the need of human observation, hence human errors are eliminated. It provides security against unwanted entry in closed monitored area without the need of other sensors like vibration or shock sensors, passive infrared detectors, motion detectors etc.

The proposed system architecture and installation is described in section II. On the other hand in section III, the algorithm and methodologies are explained. Using these methodologies, all the performed experiments and results are stated in section IV.

II. System Architecture And Installation

The paper presents the vision based security system that includes a web-camera installed in premises which can be called as a closed supervised area when no one is present in premises and security is vital for that specific period. The web camera is connected via Universal Serial Bus (USB v2.0) with the ARM11 core based target board. The camera detects the motion caused by any intruder in front of the camera. The area captured by the camera plays a substantial role to cover the whole supervised zone. The camera information is forwarded to the software system that processes the input and decides when to trigger the various events.

2.1 Hardware Design approach

Hardware architecture is based on the BCM2835 High-Definition embedded multimedia processor which consists of a 700MHz ARM11 core with VideoCore IV multimedia co-processor and an on board 512MB RAM, which has been split up into 256MB for GPU and 256MB for processing the normal operations for this application. The target board can be overclocked to 1GHz for improved performance. The overview of target board is represented in the fig. 1. The Ethernet port is useful to communicate using internet for different applications. GPIO headers are required to control the alarms and other peripherals, if needed. The expansion of root file system in Linux operating system is performed on SD card connected with target board. The RCA compliant display can be connected to see the execution and to program the target board. FM modulation libraries are ported, which modulates an alarm sound on a particular pre-set frequency in FM Band on a GPIO pin of a target board. It sets the clock generator module to be enabled and sets it to output on GPIO4 of BCM2835 processor. It also sets the frequency to 100.0 MHz, which provides a carrier. It will work from about 1 MHz upto 250 MHz by setting up a frequency in software. In our application, whenever the intruder is detected, along with other indicators buzzing alarm will be triggered on devices tuned to 100.0 MHz frequency.

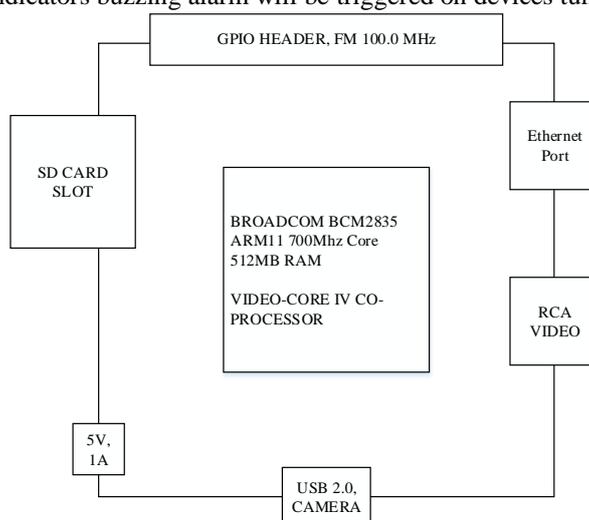


Fig. 1 Overview of target board

An Ethernet port is used to access the target board from a remote computer, which gives the access to images of the intruder captured by system at the same time. The target system can be accessed by forwarding remote authentication protocol, SSH X11.

2.2 The software design approach is segmented in to following steps.

2.2.1 Cross compilation platform and debugging

Cross compilation is a process in which compiler creates the executable code or files which can run in different platform. Compiler runs on host computer or system but produces the executable program for a target device. Initially our approach uses Microsoft Visual Studio 2012 (VS12) as a cross compiler platform which is running on windows and target architecture is ARM core based which is running on Debian Linux.

Visual studio has large collection of tools and facilities to design software applications not only for windows; it provides more flexibility for cross compilation environment. In our software design approach VS12 is used for firmware development in C language and OpenCV libraries [5] added as additional dependencies and directories.

2.2.2 Linux Debian and OpenCV porting

Debian is a Linux kernel based operating system which is one of the leading ARM based embedded development environment. It provides basic sets of programs and utilities to make embedded processor run fast and also known as light weight OS. It includes 37000+ packages for embedded and other application development.

The camera is interfaced with the target board using USB cable. The camera must be UVC (USB Video Class) compliant. UVC project delivers the essential software modules to support UVC compliant devices in Linux. UVC compliant peripherals require a common device driver and kernel support for UVC devices. OpenCV 2.4.8 package [6] is successfully ported on Linux OS, after installing the essential libraries required. The Makefiles for porting OpenCV is generated using CMake package. The debugged firmware is compiled and executed on the target board to make the system run.

III. Algorithm And Methodology

3.1 Assumptions

We assume that the monitored area is already closed and secured as well as out-of-bound of the other humans.

3.2 Algorithm

Irrespective of a human identity, recognizing a human can be usually defined as the human motion detection. The human can be determined by its certain characteristics and patterns. The more sophisticated algorithm for recognition is gait recognition or face recognition in which the human head is considered for finding the features. Our security system aims at detecting the presence of a human in the closed monitored area. The system is treated as a personal system, with an aim to track moving entity, which can be characterised by imaginable human motion and with that characteristics we decide whether it is human.

The algorithm is divided into five main tasks followed by decision block as shown in fig.2.

3.2.1 Image Acquisition

The camera module performs the task of image acquisition which provides the coloured RGB image of resolution 640X480 pixels. The ultimate objective of this task is to provide image frame through segmentation from a continuous video sequence. The camera is fixed in a static location to observe a specific area. According to the frame rates, the video sequences are generally classified into two types, higher frame rate and the lower frame rate. The higher frame rate is considered as the 30fps (frames per second) and lower frame rate is considered as 15fps. It is essential to consider the frame rates according to the hardware of the system. For motion detection, it is not necessary to acquire each and every frame; this can upsurge the calculation and creates overhead to the system. The system proposed in [1], follows the Adjacent Frame subtraction (AFS) and Background Frame Subtraction (BFS). Both the methods have its own advantages and disadvantages. In AFS due to low frame rate it creates an erroneous motion image. In BFS method, the higher frame rates cannot be used, which compromises with detection quality. To overcome these disadvantages, this system adapts a novel segmentation method called random background image subtraction.

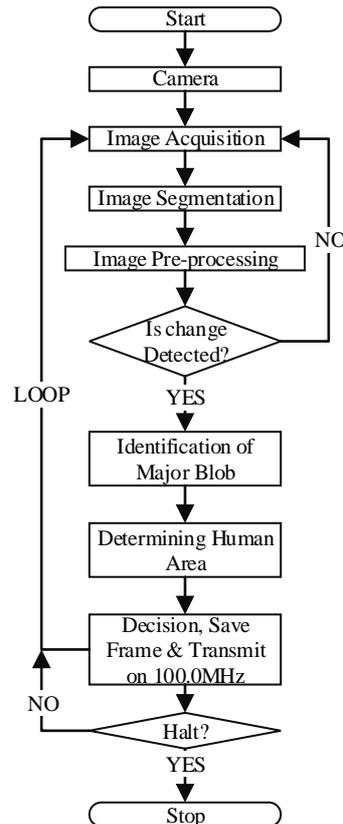


Fig.2 Algorithm Flow of Motion Detection

3.2.2 Image Segmentation

In our application we are capturing the reference image (Background Image) at a random interval and current image is captured at a defined interval. The frequent revision of background image is needed to overcome the problem of varying illumination situations inside the area. The proposed method works as follows.

Let $r(t)$ be defined as the set of time intervals, each of the values present in set provides the time to update the background image. One random number generator is casted off to provide the fractional random numbers between 1 to 5. For example $r(t)$ can be shown as stated here.

$$r(t) = \{1.054, 2.342, 3.597, 1.424, 4.643\}$$

At each initial runtime, before entering into the loop, system updates this set with the help of a random number generator. Consuming this randomised time set the background image is updated. A fractional time set is required so that the update time required for current image and referenced image do not match. The subtraction can be represented by $S(x, y, T+\tau)$ as a result of absolute image subtraction of randomly updated background image $X(a, b, r(t_i))$ and the currently captured image $C(x1, y1, T)$.

$$S(x, y, T+\tau) = \text{abs} [X(a, b, r(t_i)) - C(x1, y1, T)]$$

The subtraction image calculated at time $T + \tau$ is derived from absolute subtraction of the frames taken at random time intervals and predefined integer time interval T . Where τ is the calculation time required to take the absolute difference, $r(t_i)$ and T stands for time constants. T is always a static parameter, whereas τ and $r(t_i)$ are varying in nature; $x, y, a, b, x1, y1$ are pixel values of a particular image.

The random updating method also eliminates the Luck factor of an intruder. If the intruder enters at the same time when image is being updated, the intruder will be in the reference image, although if it happens the intruder needs to be stationary, if intruder moves it will trigger the alarm to buzz. The final absolute subtracted image is given as input to the thresholding block as described in next subsection.

3.2.2.1 Image Thresholding

This operation is a part of image segmentation process which is used to convert subtracted coloured image into binary image. In image processing sometimes we need to make a decision about pixels values of

an image or categorically reject pixels below some levels, by the use of threshold operation we can make a binary decision on image and binary image can be produced. This image can be represented as one and zero pixel values. In our application we are using Binary Thresholding operation. This operation can be described using this following relation.

$$dst(x, y) = \begin{cases} 1 & \text{IF } src(x, y) > \text{threshold value;} \\ 0 & \text{OTHERWISE;} \end{cases}$$

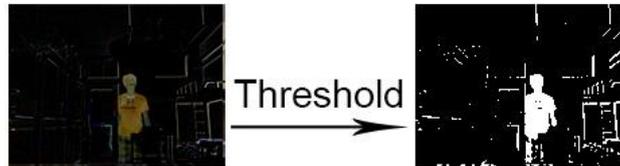


Fig.5 Image Thresholding from background subtracted image

As seen in fig. 5 it can be observed that the salt noise is present in the thresholded image. The threshold operation is performed on segmented as well as subtracted frames as described in section 3.2.2

3.2.2.2 Image Pre-processing

Morphology is a set of image processing operations to enhance the image which can be useful for this application to describe any region, shape or object characteristic from input image.

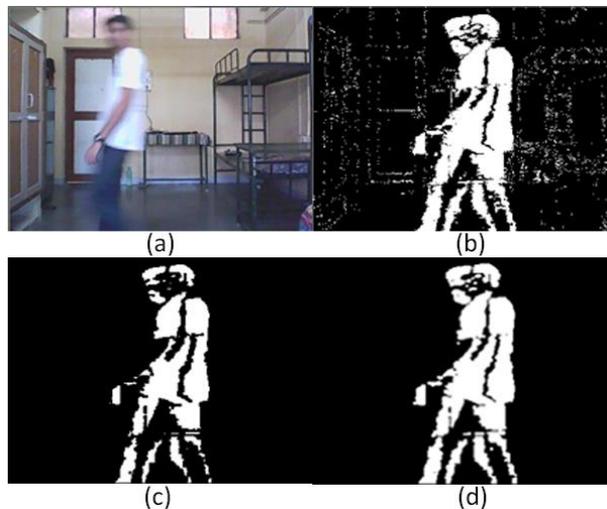


Fig.6 Current frame (a), Image Thresholding (b), Erosion Operation (c) and Dilation Operation (d)

The erosion operation is performed on a subtracted image as shown in fig. 6 (c) from thresholded image (b). The basic need of erosion operation arises due to presence of noise in the image. Because of erosion operation the background is enhanced and the thin white lines can be suppressed. Now once thin lines are subtracted the area of moving foreground object need to be enhanced back so we use the dilation operation as shown in fig. 6 (d).

3.2.2.3 Area Calculation

After selecting the major contour, the area covered by that contour is calculated and this area is compared with the results of an anthropology theory. The area from the base theory is compared by calculating the camera aperture and the area covered by camera. As we can see in fig. 7 D_{vm} stands for distant view covered by the camera in meters and the real horizontal view can be equated with the pixel value of the captured frame which is denoted as D_{vp} .

According to anthropology theory the average cross-section of anthropoid area ranges from $0.4m^2$ to $0.6m^2$. As stated in research paper by John et al. [1] the human area in pixels can be calculated by using the formula shown below.

$$\frac{A_{Hm} D_{Vp}^2}{D_{Vm}^2} = \frac{\text{meter}^2 \times \text{pixel}^2}{\text{meter}^2} = \text{pixel}^2$$

Calculations:
 $D_{Vm} = 2.5\text{m}$
 $D_{Vp} = 640\text{pixels}$

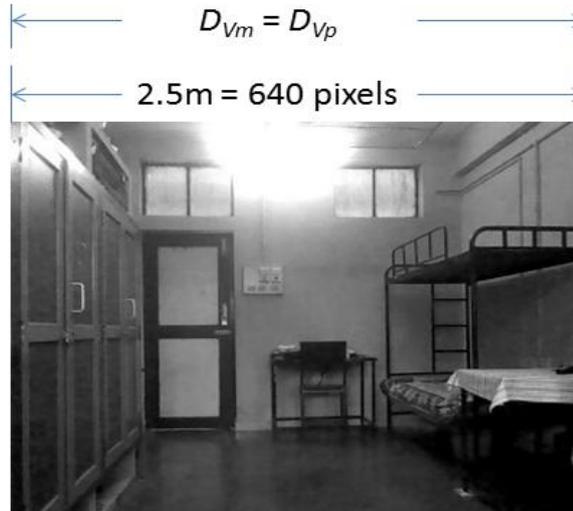


Fig.7 The area viewed from a camera

Putting number of pixels as Values,
 $A_{hp} = 26214$ [sq. pixel] for $A_{hm} = 0.4\text{m}^2$
 $A_{hp} = 39321$ [sq. pixel] for $A_{hm} = 0.6\text{m}^2$

According to this method the range of the area is set for 640X480 pixel wide image and the derived human area from the contour method can be compared with this range of theoretical values.

To avoid the erroneous warnings by the system another function is used to find the mean of the thresholded image to calculate the area of the object present in the image. This average values has a particular range against the distance, which is derived after performing practicals in the actual field. The final values derived are shown in result section IV.

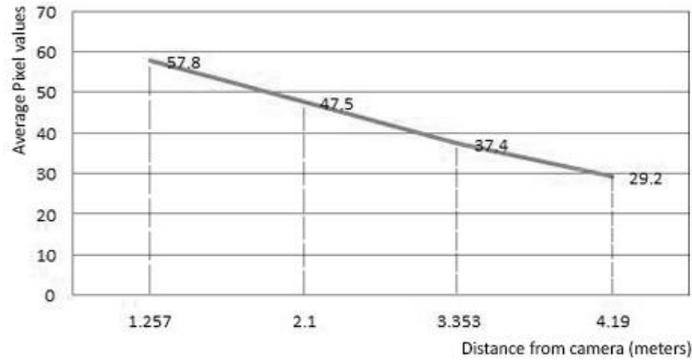
3.2.2.4 Control Action

The major objective of this paper is to determine the human motion in a closed monitored area. The control action can be generated on target hardware according to the area calculated in the previous step. The target board can be connected to internet through Ethernet port to send an image or live streaming video from area of interest to a distant user. Currently the target board is turning on an alarm light to indicate the warning caused by the intruder as well as transmits an alarm sound to the tuned FM receivers on 100 MHz frequency.

IV. Experiments And Results

The mean values are calculated by performing different experiments in the field. The intruder motion is calculated at different distances. In the simulation the intruder passes by the camera and the distance between intruder and camera is varied to calculate the different averaged values as shown in graph 1.

As it can be interpreted that the average value decreases as the distance of intruder from camera increases. These results are derived by installing camera above the height of approximately 1 meter from the ground and lens has approximately 0° deviation from the ground as shown in figure.



Graph 1 Distance vs. Average Value of thresholded image

The tests resulted from different methods are listed and explained below.

At first when there is no intruder present in the area observed by camera, meaning all objects are stationary the subtracted image is totally blank and calculated average area is exactly zero as shown in the fig. 9.

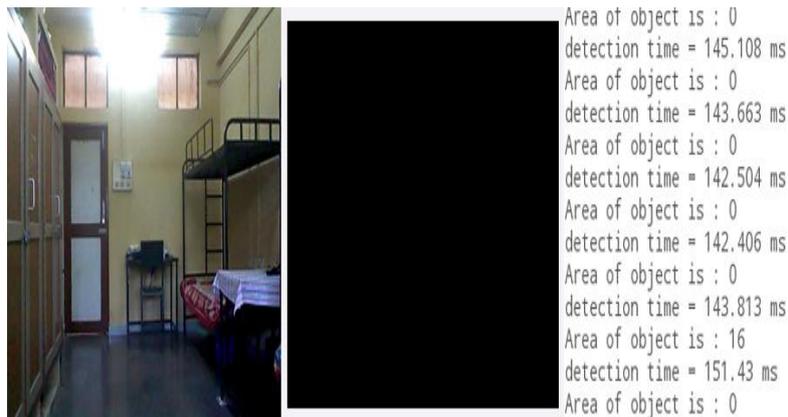


Fig.9 No intruder present in the observed area

When the intruder passes in front of camera, the white blobs are created as per the methods used in section 3.2.2.3. The fig. 10(a) shown here shows the current video captured and major blobs traced as shown in fig. 10(b). The blob areas are calculated and simulated output is shown in fig. 10(c).



Fig.10 Output screen when intruder is present

To obtain the best results the camera should be approximately 30° heading oblique to the horizontal line and 2.5m above the ground level as shown in fig 8. The camera is able to cover 4meters distance effectively. Further analysis and accuracy tables are derived using this installation and are presented in next section.

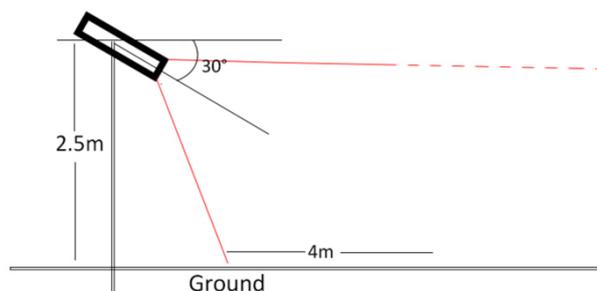


Fig.8 Camera installation for obtaining the best results

The monitoring system proposed in [1] (Pentium IV, 2.63 GHz, 504MB RAM) took 10 minutes and 16 seconds to process 1000 images i.e. 616ms to process each frame, comparatively our proposed system took approximately 181ms to process each frame i.e. 3m and 57sec to process 1000 images.

The system was kept in an ON state for 1 hour 23 minutes and a good detection rate of 86.11% was achieved. There were 72 predictable motions in which 10 were unnoticed. Non-human motions like football, moving chair were also passed in front of the camera. Though, due to less distance of object from camera or for some undetermined reasons 5 out of 20 non-human motions were detected which is as high as 25%.

V. Conclusion

The software implementation is centred on Linux kernel with porting of cross-compiled OpenCV, GPU and GUI libraries. The use of open source technologies and choosing Debian Linux as the development platform, the development cost has reduced tremendously. The algorithm flow used in this research paper makes it easy to detect a human using novel random Background subtraction method and mean area method. Hence the system is effectively able to catch the intruder moving in the supervised area. The benefits of the proposed system are low power consumption, small size, less wires, plug and play operation, and effective reduced cost. To make the system accurate, more number of camera modules should be mounted at different positions to cover the entire zone. In addition to this, in a single room, the idea of installing multiple cameras can be employed and also by enabling the communication among them via Ethernet/Wi-Fi communication, a voting scheme for a central system can be developed. The proposed scheme is practically implemented for installation in different consumer environments.

References

- [1] John See, Student Member, IEEE, and Sze-Wei Lee, Member, IEEE-” An Integrated Vision-based Architecture for Home Security System”
- [2] Donabel D. Abuan Alexander C. Abad, Jose B. Lazaro Jr., and Elmer P. Dadios, De La Salle University – “Security Systems for Remote Farm” Journal of Automation and Control Engineering Vol. 2, No. 2, June 2014.
- [3] Ren C. Luo, IEEE Fellow, Te Y. Hsu, Tung Y. Lin, National Chung Cheng University “The Development of Intelligent Home Security Robot” Proceedings of the 2005 IEEE International Conference on Mcchatronics July 10-12,2005, Taipei, Taiwan.
- [4] QusayIdreesSarhan, Department of Computer Sciences, University of Duhok, Duhok, Iraq. “An Integrated Real-Time Vision Based Home Security System” International Journal of Advancements in Research & Technology, Volume 2, Issue 5, M ay-2013.
- [5] Gary Bradski and Adrian Kaehler, LearningOpenCV(O'REILLY Publications)
- [6] [Online] “OpenCV Reference Manual v2.1”at www.opencv.org