EAI Endorsed Transactions
on Scalable Information Systems

Research Article EAI.EU

Intelligent Virtual Security System using Attention Mechanism

[a]Wasim Ahmad Khan, [b]Hafiz Usman Akmal, [c]Ahmad Ullah, [d]Aqdas Malik, [e]Sagheer Abbas , [f]Abdullah Ahmad, and [g]Abdullah Farooq

[a,e]National College of Business Administration and Economics Lahore, Pakistan
[b,c,d,f,g]Punjab University College of Information and Technology Lahore, Pakistan

Abstract

Within the past few years, organizational security has emerged as a critical challenge for industries, security agencies, academic institutions and organizations. The current security system lacks efficiency which influences organizations' security parameters negatively. While dealing with a violent or terroristic activity human guards may not be able to response effectively under pressure, due to some limitation or lack of continuous attention. In present conditions where guards are not active enough to generate a quick response against threat, human eagerly need to design an effective virtual security system which would be responsible to analyze collective measures, process, and procedure that ensure protections of virtualization environment. In order to increase efficiency of security, eliminate security holes and to obtain quick response towards suspicious activities, there is need to replace humansecurity-guard with a virtual-security-guard system. This virtual guard will be an embedded machine, which will act like a complete security guard that will work under 'Human Based Psychology'. This system would be reflection of an ideal virtual security guard which will generate intelligent and quick response towards abnormal activities by making intelligent decisions based on selective attention mechanism. The proposed system can be efficient enough to focus both functional and non-functional requirements (resistibility in smoke, fog, rain, dark and dust).

Keywords: Virtual Security, Artificial Intelligence, Computer Vision, Attention, Motion Detection, Blob Detection.

Received on 21 December 2017, accepted on 03 January 2018, published on 13 April 2018

doi: 10.4108/eai.13-4-2018.154473

1. Introduction

The business organizations, institutions and military agencies of 21st century are highly dependent on effective technology. These systems are quite insecure due to internal and external sources. Currently implemented security parameters are not rational enough to perform intelligently in complex situations. The conventional security methods should be replaced by virtual security systems. The current security systems can be transformed in to advance level security system by implementing latest technologies. In this era, security agencies themselves are most insecure departments as compared to other organizations [1]. In order to make security agencies secure, technologists eagerly need to implement a virtual security guard system. This system should be based on human based psychology of a human guard. It should implement the features and actions of a human guard, the way guard observes things from his surroundings, capture the surroundings, distinguish selective data from multi-party interaction [2]. The security hole present in our organizations can be filled using effective computer vision techniques.

The most common surveillance systems available in our surroundings are based on simple video recording feature. Cameras with additional features like motion detection, face detection, people count are available in the market but are not implemented as a single unit for complex organizations. It is required to set in all these technologies together in one unit to compose a strong system.

These real-time virtual security systems in parallel to video capturing/recording, targets on making analysis on video under set of rules & formulas. The input data is processed within milliseconds, the set of rules and algorithms embedded on these systems observe different behaviors & patterns of motion [3]. Right after the processing phase, immediately a response is generated by these systems that is based on human based psychology. The security alerts or responses generated by automate systems are installed according to contingency plans of organizations. The proposed systems has minimized manual human monitoring to some extent. However, a lot of efforts are required to implement an efficient automate virtual security system.
A closer view indicates that the most important factors on which these automate surveillance system are based on are motion detection, attention and blob detection.

1.1Selective Attention

The system can be only made intelligent, if it focuses on "What is required?" It must focus on selective objects by ignoring other factors as noise. The system will be placed in different environments under different security circumstances. Selective Attention refers to information processing under the defined context [4]. In this mechanism, the function receives data, processes it, and filters the prioritized data. Top down and Bottom up are two approaches to optimize detection speed of attention in any scenario [5].

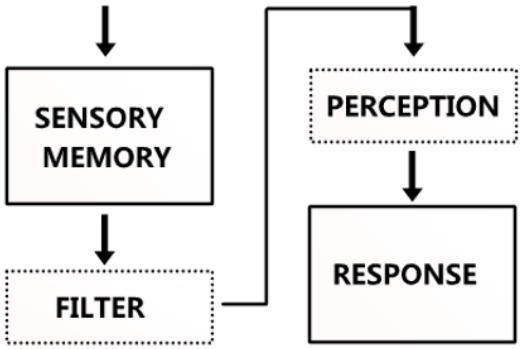

Figure 1 - Filter Model for Selective Attention

1.2Motion Detection

In order to develop an efficient surveillance system, motion detection has a very significant role. Motion detection is based on monitoring algorithms which are formulated to track the motions or detect activities. Motion detection extends into an advance level when it is distinguished in to different type of motions based on species and environment. In most common detection algorithm 'frames comparison' is performed, changes are highlighted and if required, action is performed.

1.3Blob Detection

In order to implement hard and fast rules on received data and to distinguish objects from frames; Blog detection is performed. Blob Detection methods identifies objects on basis of difference in their properties .i.e. Color, Brightness and difference in neighboring pixels.

1.4 Virtual Security

Virtual Security system refers to process of analyzing input from algorithms and generating quick response(s) based on human psychology. Human based psychology means that the system interprets any scenario from environment like a human. The information from environment is gathered in form of frames and selective attention features are extracted. A signal is created which is passed to short term memory. Based on the scope and target, if the signal has some importance then it is forwarded to activation mechanism that generates proper response. Here, proper response means action toward detected problem.
The trend of virtual surveillance system is getting viral day by day. Organizations are looking for efficient security parameters to fill the security flaws. Systems are deployed in institutes according to the environment and level of security.
Well, the current embedded systems are not efficient enough to smartly distinguish between real threats and false assumptions. The future demands use of advance technologies bonded with precise algorithms to generate a powerful protection system.

This is the body text with no indent. This is the body text with no indent. This is the body text with no indent. This is the body text with no indent. This is the body text with no indent. This is the body text with no indent. This is the body text with no indent. This is the body text with no indent. This is the body text with no indent. This is the body text with no indent. This is the body text with no indent. This is the body text with no indent. This is the body text with no indent.

2 Related Work

The virtual security system has been presented in different editions, in past. Number of API's (Application Program Interface) have been defined and implemented by researchers and developers. All these proposed works vary from each other, some have plus point but most of them are incomplete and lack full functionality. Let's begin from the end of 20th Century.

2.1 Visual Tracking using Closed-Words

In the late 90's, concept of "Close World Track" was introduced for visual tracking. It is basically tracking in space time regions in which all pixel are detected from some scenes. With help of Video Annotation, each Frame

is analyzed in player's movement. The collision generates colors that detect blob [6].

Figure 2 - Football Player Motion Detection

The missing feature in this system was that there was no tracking of motion rate of football players.

2.2 Cognitive Active Vision for Human Identification

This system is based on inference Engine that extract information from captured scene. Pan Tilt Zoom Camera is used to study scene, analyze it and do Facial Recognition of People in scene with the help of applied formula [7].

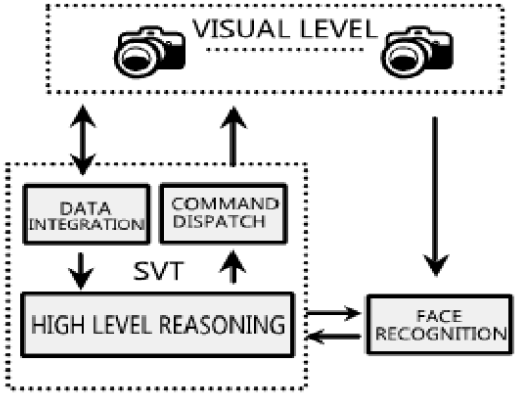

Figure 3 - Work Flow of Cognitive Active Vision

Details of System is included in distributed Camera System which had:

- Two cameras (Static Cameras and PTZ (Pan Tilt Zoom Cameras)
- SVT - Supervisor Process.
- The command Table gives command on which PTZ act.
- The Image Table contains captured images from scenes which are parsed for Face Recognition.

On-Line Inference and Camera Control following activities were performed:

- Tracking target and acquire image of face Performed by:
- Situation Graph Trees
- High Level Reasoning.

2.3 An Attention Model for Extracting Components that merits Identification

In this attention model, the blobs were extracted by sorting out the meaningful information from the images. The blobs were detected by the following steps:

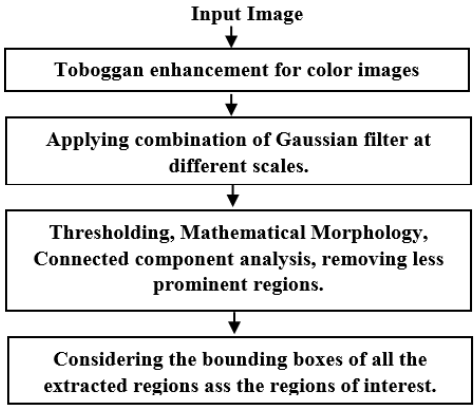

Figure 4 - Blob Detection Steps

The system was implemented to remove the least significant regions from the image under observation [8].

The regions of least interest are removed by following steps:

- Rank overlapping regions.
- Select & compute the prominence of each region.
- Remove the region with low prominence.
- Repeat.

The formula used for computing prominence factor was:

$$Prom(\omega_i) = \frac{\sum_k R(k)}{Area(\omega_i)} \quad \textbf{(1)}$$

Using bottom up approach, the accuracy for blog detection of this algorithm was 79.8% [9].

2.4 Motion Detection Surveillance System Using Background Subtraction Algorithm

This paper provides motion detection with the help of robustness algorithm and detect motion rate using technique of background subtraction. It captures static image and whenever it detects any change, it subtracts current image from previously captured image to analyze difference [10]. It gives object blobs which are compared by last image frame in which object was detected. In this way, motion is detected from background.

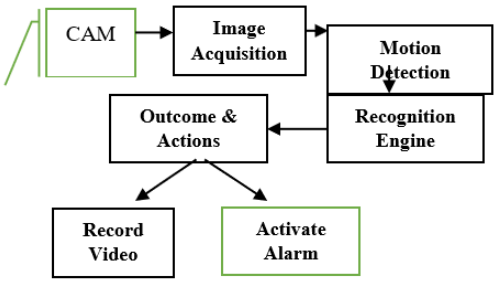

Figure 5 – Basic Blob Detection

Mathematical Equation:
P[F(t)] = P[I(t)] – P[B] (1)

Where, F refers to foreground resulted image after computation at time t.

3. Comparative Analysis:

Paper	Basic Principal	Features
Visual Tracking using Closed-Words	Visual Tracking	Blob Detection
Cognitive Active Vision for Human Identification	Facial Recognition	Face Detection
An Attention Model for Extracting Components that merits Identification	Attention	Blob Detection
Motion Detection Surveillance System Using Background Subtraction Algorithm	Human Motion Detection	Blob & Motion Detection
Visual Security System using Attention Mechanism	Attention	Motion Detection, Face Detection, Eye Detection, White Light Detection, White Light Blink Detection

This analysis is comparison of features and accuracy between related works and the proposed system. Previously composed systems were based on single principle .i.e. Visual Tracking [11], Attention, Human motion but currently proposed system is applicable to focus on all aspects, as it follows “One system, maximum features”.

4. Proposed System

This paper proposes a method of visual attention [4] recognition using cognitive agents. It uses different techniques to generate visual attention i.e. Face detection, Eye detection, Blob detection and Motion Rate Detection. To start with basic and initial steps that helps to generate attention, some important features are introduced. These utilized features include motion rate detection, blob detection, white light blob detection, face and eye detection of cognitive agents, brightness values detection from environment and fluctuates attention on detection of all these features.

Proposed virtual security system uses a simple camera and performs different detection algorithms on video stream and generates a suitable attention as a result of all detection. It takes video stream and divide video in frames of image to apply detections on each frame. Face Detection is done by using haarcasade technique on each frame of image. It analyze each frame and count number of faces after face detection method. Profile face detection method is also implemented in this system. It helps to detect face of a person whose face is not exactly in front of camera or standing in a little bit tilt position. Eye Detection is also performed by using haarcasade technique which helps to count number of eyes in image frame of video stream. Motion Rate is detected by analysis of current frame value of scene with previous frame values and then take average of all computed frames which gives differentiation of frames with other on each timer tick. Blob Detection is basically a White Light Blob Detection which takes a bitmap image and perform grey scale conversion to get gray image. Background subtraction is then performed on gray image to extract light features and detect light blobs from image. Both, Motion Rate and Light Blob Detection increase attention factor according to environment. Blinking is also an important feature which is then achieved by Light Blobs and maximize attention level on Light Blink (On/Off Loop). Light blinking factor is considered as a suspicious activity in proposed virtual security system because it shows hindrance of view of camera and due to this activity, system generate some response in form of attention. Each detection value of number of faces, number of eyes, motion rate value and brightness value of each frame has suitable impact on attention factor which is then visualized in form of graph with some parameters.
Proposed system is implemented in different phases, in first phase, attention is visualized by number of faces. In second phase, attention is generated by combination of number of faces and eyes. In third phase, motion rate factor added in system along with face and eye detection method. In forth phase, attention is generated on graph with light blob detection with all three detections. Work flow model of proposed system from start of input stimuli to some response of system as attention is shown in Figure 6.

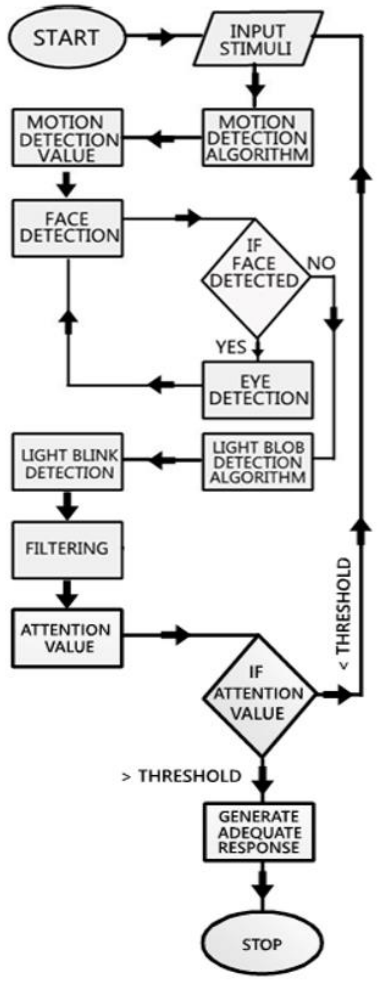

Figure 6 – Proposed Algorithm Flow Chart

Scenario:

Consider an institution in which proposed mechanism is installed at the entrance of hall. The purpose of this system will be to figure out the attention rate by analyzing the movement on the basis of people count. The attention rate will distinguish according to number of people and motion in the scene.

The step-by-step flow for the composed system is as follows:

- Video camera installed on the site acts as input stimuli. It will capture the scene as video frames.
- The frames are forwarded to Motion Detection Algorithm. Algorithm will compare current frame with previous frame to compute motion detection value.
- After calculating the motion rate, the target of system will be to detect Face.
- **IF,** face will be detected by face detection algorithm, it will detect eye. The algorithm will keep detecting the face and eyes until no more face is found. The result will give value of people count.
- **ELSE,** when no more face will be detected. The algorithm will process the data through Light blob Algorithm. The image frames will pass through gray scale conversion, and will extract light factors from the grayscale image. This is how image detection and blinking will be observed.
- The purpose of getting light blinking values is to point out, if some object in frame is trying to block the view of camera on sight.
- Once the above steps are completed. **Data** is refined by filtering & removing noise factors.
- The final output extracted from this mechanism will be **Attention Rate.** This value will be in between the range of 0 and 1.
- **IF,** attention rate value will be greater than the defined Threshold Value .i.e. T, **THEN** system will generate an adequate response .i.e. Alert or Emergency call.
  - **ELSE,** the system will keep working on calculating selective attention until it is paused/stopped.

4. Conclusion and Future Work

Accuracy of Visual Attention requires proper information that exist in environment. The surveillance systems established in past opened gates of research & development in context to security systems but most of them are incomplete and lack ability to deal with all functional and non-functional requirements. The implemented automated Virtual Security System is an innovation that will re-direct developers to compose an ideal virtual surveillance system. Number of features implemented in this project will prove helpful to enhance security using attention mechanism. By embedding this project in society, security system will be replaced by virtual security system that will overcome responsibilities of guards. In this way, death rates of human security guard will decrease. In future, fuzzy logic and neural network techniques will be implemented in this system which will generate attention output with the help of different parameters and rules. Fuzzy logic will provide precise and accurate value of attention on the basis of face and eye count factors. Moreover, the efficiency of guards' response time will fill the security hole and much more. The only target of this project will not be limited to security of guards but it will be intended to reduce expenditure on security. More factors will be analyzed in the next version to get more accurate attention value, in order to implement human like attention mechanism.

5.References

1. Erwin, S. I., Magnuson, S., Parsons, D., & Tadjdeh, Y. (2012). Top five threats to national security in the coming decade. *National Defense*, *97*(708), 30-37.
2. Stefanov, K., & Beskow, J. (2016). A Multi-party Multi-modal Dataset for Focus of Visual Attention in Human-human and Human-robot Interaction. In *LREC*.
3. Reese, H. (2016). How one AI security system combines humans and machine learning to detect cyber threats.[Online].Available:http://www.techrepublic.com/article/how-one-ai-security-system-combines-humans-and-machine-learning-to-detect-cyberthreats/.
4. Remington, A., Swettenham, J., Campbell, R., & Coleman, M. (2009). Selective attention and perceptual load in autism spectrum disorder. *Psychological Science*, *20*(11), 1388-1393.
5. W. Khan, M. Khan and D. Hussain. (2016) "Integrated approach of human-robot interaction on basis of top-down and bottom-up attention approach using visual and verbal affective intent in agents", 4(2).
6. Intille, S. S., & Bobick, A. F. (1994). *Visual tracking using closed-worlds*. Media Laboratory, Massachusetts Institute of Technology.
7. Utsumi, Y., Sommerlade, E., Bellotto, N., & Reid, I. (2012, May). Cognitive active vision for human identification. In *Robotics and Automation (ICRA), 2012 IEEE International Conference on* (pp. 1238-1245). IEEE.
8. Serre, T., Wolf, L., Bileschi, S., Riesenhuber, M., & Poggio, T. (2007). Robust object recognition with cortex-like mechanisms. *IEEE transactions on pattern analysis and machine intelligence*, *29*(3), 411-426.
9. Jahangiri, M., & Petrou, M. (2009, November). An attention model for extracting components that merit identification. In *Image Processing (ICIP), 2009 16th IEEE International Conference on* (pp. 965-968). IEEE.
10. Sahu, A. K., & Choubey, A. (2013). Motion detection surveillance system using background subtraction algorithm. *International Journal of Advance Research in Computer Science and Management Studies*, *1*(6).
11. "Visual Tracking and Academic Performance – Part 1", *The VisionHelp Blog*,2017.[Online].Available:https://visionhelp.wordpress.com/2010/11/27/visual-tracking-and-academic-performance-part-1/. [Accessed: 31- May- 2017].
12. Erwin, S. I., Magnuson, S., Parsons, D., & Tadjdeh, Y. (2012). Top five threats to national security in the coming decade. *National Defense*, *97*(708), 30-37.