

An Event-Driven Multiple Objects Surveillance System

Preliminary Communication

Saeed Mina Qaisar

Effat University, Jeddah, KSA
College of Engineering, Department of Electrical and Computer Engineering
Jeddah, Saudi Arabia
sqaisar@effatuniversity.edu.sa

Dija Sidiya

Effat University, Jeddah, KSA
College of Engineering, Department of Electrical and Computer Engineering
Jeddah, Saudi Arabia

Mohammad Akbar

Effat University, Jeddah, KSA
College of Engineering, Department of Electrical and Computer Engineering
Jeddah, Saudi Arabia

Abdulhamit Subasi

Effat University, Jeddah, KSA
College of Engineering, Department of Information Systems
Jeddah, Saudi Arabia

Abstract – Traditional surveillance systems are constrained because of a fixed and preset pattern of monitoring. It can reduce the reliability of the system and cause an increased generation of false alarms. It results in an increased processing activity of the system, which causes an augmented consumption of system resources and power. Within this framework, a human surveillance system is proposed based on the event-driven awakening and self-organization principle. The proposed system overcomes these downsides up to a certain level. It is achieved by intelligently merging an assembly of sensors with two cameras, actuators, a lighting module and cost-effective embedded processors. With the exception of low-power event detectors, all other system modules remain in the sleep mode. These modules are activated only upon detection of an event and as a function of the sensing environment condition. It reduces power consumption and processing activity of the proposed system. An effective combination of a sensor assembly and a robust classifier suppresses generation of false alarms and improves system reliability. An experimental setup is realized in order to verify the functionality of the proposed system. Results confirm proper functionality of the implemented system. A 62.3-fold system memory utilization and bandwidth consumption reduction compared to traditional counterparts is achieved, i.e. a result of the proposed system self-organization and event-driven awakening features. It confirms that the proposed system outperforms its classical counterparts in terms of processing activity, power consumption and usage of resources.

Keywords – Event-driven awakening, embedded processing, face detection, resources utilization, surveillance

1. INTRODUCTION

Recent technological advancements have revolutionized the usage of surveillance. It is used in a variety of applications like detection of abnormal behavior, protection of property, public safety, crime prevention, parking lots, railway line crossing, monitoring of operations, sports, monitoring homes of elderly people, parks, traffic control, warehouses, green houses, prohibited areas, etc. [1]–[12].

Traditional surveillance systems are constrained because of a preset pattern of monitoring. It can result in the generation of false alarms [45], [46], and thus cause a reduction in the reliability of the system [13]–[15]. Furthermore, irrespective of the intended object presence, the system continually remains active [13]–[15]. Therefore, it causes a useless increase in the system processing activity and results in increased consumption of power and utilization of resources by the system [16].

Within this framework, a novel human surveillance system is proposed. It overcomes these downsides of traditional counterparts by intelligently exploiting and integrating event-driven awakening and self-organization features. It is capable of effectively monitoring multiple objects present within the surveillance system range. Works presented in [16]–[19] are sources of inspiration to integrate the event-driven awakening feature into the proposed system. The proposed solution also possesses an interesting feature of self-organization [20], [36]. Its pattern of monitoring and processing is adapted according to activities of humans within the system range. It makes the devised system perform better in comparison to traditional ones in terms of processing activity and reliability. Here, the term processing activity refers to a direct relationship with the utilization of resources and consumption of power. A less processing activity causes a reduced utilization of resources like storage memory, transmission bandwidth, power consumption, etc.

The system employs a smart combination of a sensor assembly and a robust classifier [32], [33]. It suppresses the generation of false alarms and improves system reliability. These interesting features are achieved by smartly integrating event-driven awakening and its self-organization features into the system.

Section 2 describes the devised system principle along with its different modules. The field of the view coverage principle is presented in Section 3. Section 4 discusses the functionality of the proposed system with the help of Algorithmic State Machine (ASM) charts. Section 5 describes system prototype implementation and post-implementation testing results. Finally, Section 6 concludes the article.

2. THE PROPOSED SYSTEM PRINCIPLE

The principle of the proposed system is shown in the block diagram given in Figure 1. It shows that the proposed system front-end is based on two similar modules. The modular design approach simplifies system implementation and debugging by controlling the design and implementation complexity. Different system blocks are described in the following subsections.

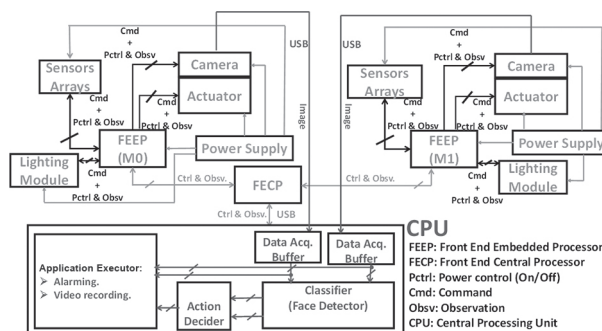


Fig. 1. The proposed surveillance system block diagram

2.1. SENSOR ARRAYS

2.1.1. Event detectors

The proposed system is designed for human surveillance. Humans belong to the class of living objects [21] that emit infrared rays [22]. Event detectors always remain in the active mode. They snoop the presence of a living object by revealing infrared rays within the surveillance range [23]. Once an event is sensed, the information is transmitted to the front end embedded processor (FEEP) and the front end central processor (FECF).

2.1.2. Object location detectors

This module is activated by the FEEP according to the event detector output. After activation, it snoops living object locations in the system range [24]. Later on, camera orientations are managed by FEEPs as a function of the location detector output. According to the targeted application, the deployment of location detectors within the system range can be realized in different geometrical configurations. Depending on the output of this module, camera orientations are determined by FEEPs.

2.2. FRONT END EMBEDDED PROCESSOR (FEEP)

When an event occurs, firstly the corresponding FEEP and the FECF are activated. Later on, in coordination with the FECF, the FEEP manages system modes of operation, status observation and control of different system components like the camera, the actuator, sensor arrays and a lightning module.

2.3. CAMERA AND ACTUATOR

Each module contains one camera. It is mounted on an actuator and its orientation can be updated by the FEEP [25], [26]. The actuation commands are generated by the FEEP according to location detector output. It pilots the actuator in the desired pattern and orients the camera towards the targeted objects. The procedure is further discussed in Section 3.

2.4. LIGHTING MODULE

Appropriate luminance required for the smooth functioning of the system is provided by this module [27], [28]. The activation of the lighting module is controlled by the FEEP as a function of the event detector and luminosity sensors integrated in the lighting module outputs [29].

2.5. FRONT END CENTRAL PROCESSOR (FECF)

It manages both front-end modules and acts as a bridge between them and the Central Processing Unit (CPU). It receives information about system status via FEEPs through Inter-Integrated Circuit (I²C) ports [30], and the CPU through the Universal Serial Bus (USB) port [31]. Both FEEPs collect this information through their respective sensor arrays. The FECF employs this information to manage the system functional states (cf. Section 4).

2.6. CENTRAL PROCESSING UNIT (CPU)

This module is connected to the FECP and cameras via USB ports [31]. It coordinates with the FECP in order to update system states. Moreover, after activation, it receives images from cameras. It is composed of two data acquisition buffers, one classifier for face detection, one action decider and one action executor. When the event occurs, the CPU is triggered by the FECP. After activation, the CPU manages the acquisition of images from (an) active camera(s). Later on, these images are analyzed for the purpose of human face detection.

It acquires images from (an) active camera(s) and analyzes them for human face detection. The face detector block is based on the Viola Jones Algorithm [32]–[34]. It provides two outputs for each front-end module, respectively.

The action decider employs outputs of the face detector block. It remains standby along with the action executor, except if one or both face detector outputs are high. The presence of one or more humans is confirmed by a high logic output of the face detector block. It activates the application executor block, which executes operations in parallel for both front-end modules. It can generate alarms and record images and video clips of the intended period [35]. This module is disengaged when outputs of both face detectors become low.

3. FIELD OF VIEW COVERAGE

In [28] and [37], the surveillance system angle object coverage problem is discussed. In the proposed approach, the overall coverage is 360°. It is split into different zones of localization. Each location sensor can conceal a specific portion of the complete field of view. In [28] and [37], authors employed multiple stationary cameras in order to cover the overall field of view. However, the suggested solution is composed of only two rotatable cameras. Location sensor output is employed to pilot the actuators. It is done in order to orient cameras towards the active zones. In this way, reliable and real-time surveillance of all active zones is ensured. The procedure is described with the help of Figure 2, where O presents the system origin and L is a horizontal axis.

Figure 2 shows that the system employs two cameras C_1 and C_2 . They are placed on actuators A_1 and A_2 and are responsible for covering the horizontal angles between $[0^\circ; 180^\circ]$ and between $[180^\circ; 360^\circ]$, respectively. By default, C_1 and C_2 remain in their home positions and cover Z_{p_2} and $Z_{p_{N-1}}$, respectively. Location sensor outputs are employed to update locations of C_1 and C_2 towards other active zones.

The system plane coverage view is defined by inequality (1).

$$0 < \frac{1}{2} r^2 \theta < \pi R^2, \quad (1)$$

where r is the radius from the origin to the target point and R is the devised system overall coverage radius.

The system also contains a set of event detectors $SE = \{E_i, \text{where } i \in (1, 2, 3, \dots, M)\}$. Their locations are prearranged.

An event detector location is given by (r_{ei}, θ_{ei}) . Here, r_{ei} is the i th event detector distance from the origin. θ_{ei} represents the horizontal placement angle of the i th event detector. The i th event detector sensing zone, Z_{ei} , is given by equation (2).

$$Z_{ei} = \frac{1}{2} (r_{ei})^2 \alpha_{ei}. \quad (2)$$

In equation 2, α_{ei} is the central angle of the i th event detector with respect to its position. It is a function of physical properties and the fabrication process of the employed sensor. r_{ei} is a radial position of the i th event detector from the origin O . It lies within the range of $0 < r_{ei} < R$, where R is the maximum range of the plane coverage zone from the origin O .

The event detection field of view can be divided in a uniform or non-uniform way. It can be done by employing event detectors of different types. The choice should be made according to the sensitivity of different zones in the intended application [16]. A case of uniform distribution of the event detection field, among different event detectors $Z_{e1}, Z_{e2}, \dots, Z_{eM}$, is shown in Figure 2. In this case, a unique type of event detectors should be employed in the system.

A set of location sensors, $SP = \{P_j, \text{where } j \in (1, 2, 3, \dots, N)\}$, is also integrated in each front-end module of the suggested system. Their locations are prearranged. The location of a position sensor is given by (r_{pj}, θ_{pj}) . Here, r_{pj} is the j th positioning sensor distance from the origin. θ_{pj} represents the horizontal placement angle of the j th positioning sensor. The j th positioning sensor coverage zone, Z_{pj} , is defined by equation 3.

$$Z_{pj} = \frac{1}{2} (r_{pj})^2 \alpha_{pj} \quad (3)$$

In equation 3, α_{pj} is the central angle of the j th location sensor with respect to its position. It is a function of physical properties and the fabrication process of the employed positioning sensor. r_{pj} is the radial position of the j th location sensor from the origin O . It lies within the range of $0 < r_{pj} < R$, where R is the maximum range of the plane coverage zone from the origin O .

Event and location detectors can have different sensing radii. However, a case where $r_{ei} = r_{pj} = R - r_{ei} = R - r_{pj}$ is shown in Figure 2. Here, R is the overall coverage radius of the system.

A set of multiple objects, $SH = \{H_a, \text{where } a \in (1, 2, 3, \dots, k)\}$, can be covered by the proposed system. Locations of intended objects are not static, within the range of surveillance. They are sensed and resolute in a real-time fashion with the help of event detectors, location sensors and FEEPs.

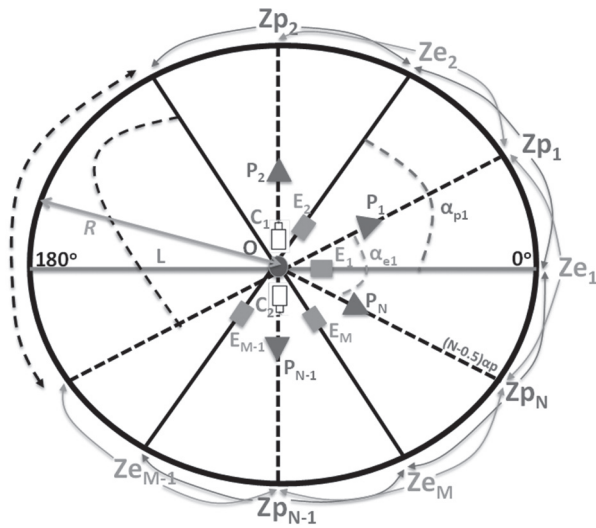


Fig. 2. The coverage principle of the proposed system

4. SYSTEM FUNCTIONALITY

The functionality of the proposed system is described by means of an ASM chart displayed in Figure 3. To improve power efficiency of the system, after initialization all system modules remain in the standby mode except for event detectors. Cameras mounted on actuators remain at their respective home positions.

The current system state becomes not equal to *STANDBY* on the recognition of an event. It triggers *FEEPs* and the *FECp*. Later on, *FEEPs* and the *FECp* manage the activation of other system blocks. In coordination with *FEECPs* and *CPU*, the *FECp* monitors the status of location sensors, face detector, cameras, actuators and the lighting blocks. Later on, the *FECp* updates functional states of the system according to these observations. Different system states are discussed next.

4.1. STANDBY STATE

In this state, except for low-power event detection sensors, all other system modules are kept in the standby mode. It assures a better utilization of system resources. The system functions and processes only relevant information and conserves its assets. Upon event detection by one of the front-end event detector array, the state of the system is updated to *SNOOPING*.

4.2. SNOOPING STATE

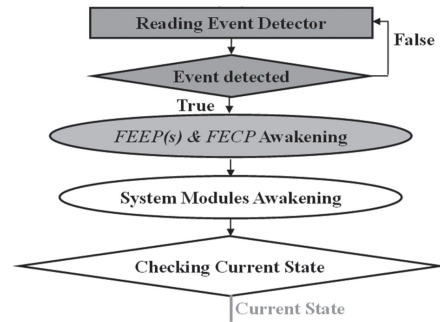
In this case, in coordination with *FEEPs* and the *CPU*, the *FECp* keeps observing the status of both front-end location sensors. This state is split into three different cases as a function of location sensor output.

Case 1: If no location sensors have found the intended object, then the state of the system is modified to *STANDBY_MODE_ACTIVATION*.

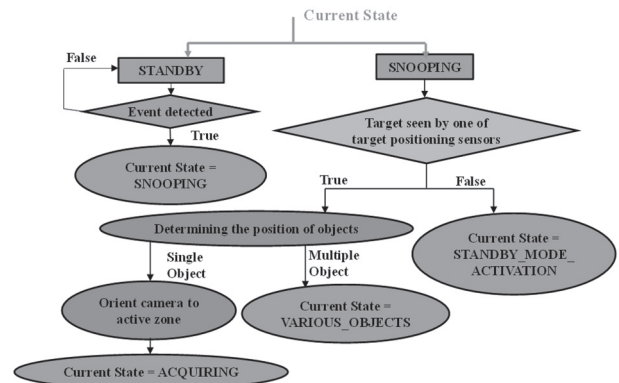
Case 2: If one location sensor has sensed the intended object, then the location of the target is first decided by the concerned *FEEp(s)*. It is attained by us-

ing the target sensor predefined location table. Later on, the concerned camera(s) is oriented towards the respective active zone. Finally, the state of the system is updated to *ACQUIRING*.

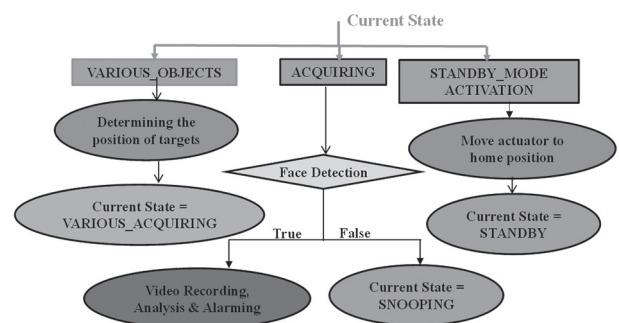
Case 3: If more than one location sensors have sensed the intended objects among one or both front end sensor arrays, then the state of the system is modified to *VARIOUS_OBJECTS*.



(a)



(b)



(c)

Fig. 4. The *VARIOUS_ACQUIRING* ASM

In this state, images are captured and analyzed for face detection. When at least one face is recognized, alarms are generated and videos are recorded for the corresponding cameras for a predefined period. Later on, the state of the system is modified to *SNOOPING*. On the other hand, if no face is recognized on the acquired image, then the state of the system deprived of alarming and recording of a video is modified to *SNOOPING*.

4.6. STANDBY_MODE_ACTIVATION STATE

Cameras are oriented towards their respective home positions in this state. Later on, the state of the system is modified to *STANDBY*.

5. IMPLEMENTATION AND RESULTS

A system prototype is realized in order to exhibit the appealing features of the proposed system. Commercially available appropriate components are employed to achieve an expeditious prototype development. System costs, size and power consumption could be further optimized by employing a specific system design and development in place of the standard evaluation card based implementation.

The desired field of view of the developed system is between $[0^\circ; 360^\circ]$. The surveillance field is divided into six different positioning zones (cf. Figure 5). It is done in accordance with the employed cameras and positioning sensor characteristics, [16], [42]. It allows to properly cover the system overall field of view between $[0^\circ; 360^\circ]$. The first front-end module, M_o , covers the surveillance field horizontal angle between $[0^\circ; 180^\circ]$. It is composed of zones $[Z_{p1}, Z_{p2}$ and $Z_{p3}]$. The second front-end module, M_r , covers the remaining surveillance field horizontal angle between $[180^\circ; 360^\circ]$. It is composed of zones $[Z_{p4}, Z_{p5}$ and $Z_{p6}]$. The overall system coverage radius $R=4m$ is selected. In this case, the system coverage view can be defined by inequality 4.

$$0 < \frac{1}{2}r^2\theta < 16\pi, \quad (4)$$

where r is the radius from the origin to the targeted object and $R=4m$ is the overall coverage radius of the developed system.

The *FECP* and *FEEPs* are implemented with Arduino UNO [38], [39]. A specific C-based development is realized for the front-end embedded blocks.

The lightning block is implemented with power efficient light-emitting diodes (*LEDs*) and light dependent resistors (*LDRs*) [16]. The *CPU* is implemented on a personal computer (*PC*). The classifier, application decider and application executor blocks are implemented with MATLAB. These blocks run on the *CPU*. The liaison between the *FECP* and *FEEPs* is achieved via I²C ports. The *CPU* is connected to the *FECP* via the UART-USB bridge [47]. The images are transferred from cameras to the *CPU* through USB ports.

Figure 5 shows that the realized system is composed of two cameras C_1 and C_2 mounted on actuators A_1 and A_2 , respectively. LifeCam HD-3000 with a central angle of 70° are employed as cameras [16], [40], and S3003 servo motors are employed as actuators. In order to orient camera lenses towards the active zones, actuators are piloted by *FEEPs* as a function of sensor array output. Rotation angles of actuators are adjusted in such a way that the C_1 lens remains at 30° , 90° and 150°

for Z_{p1} , Z_{p2} and Z_{p3} , respectively. In the same way, the C_2 lens remains at 210° , 270° and 330° for Z_{p4} , Z_{p5} and Z_{p6} , respectively.

The system also employs two event detectors, E_1 and E_2 . These are implemented with passive infrared (PIR) sensors and have a central angle of 180° , [16], [41]. E_1 covers the system detection field between $[0^\circ; 180^\circ]$, and E_2 covers the system detection field between $[180^\circ; 360^\circ]$. E_1 and E_2 locations are predetermined at $(r_{e1}=0.1cm, \theta_{e1}=90^\circ)$ and $(r_{e2}=0.1cm, \theta_{e2}=270^\circ)$, respectively. They have equal sensing radii of $r_{e1}=r_{e2}=(4-0.1)m$, along with equal respective central angles of $\alpha_{e1}=\alpha_{e2}=180^\circ$ (cf. Figure 5). It results in two equal event detection sensing zones:

$$Z_{e1} = Z_{e2} = \frac{1}{2}(3.9)^2 \cdot \pi. \quad (5)$$

The specific coverage zone for each event detector is given by equations 6 and 7, respectively:

$$Z_{e1} = \frac{1}{2}(3.9)^2 \cdot \theta, \text{ where } 0 \leq \theta \leq \pi, \quad (6)$$

$$Z_{e2} = \frac{1}{2}(3.9)^2 \cdot \theta, \text{ where } \pi \leq \theta \leq 2\pi. \quad (7)$$

A set of six positioning sensors, $SP = [P_1, P_2, P_3, P_4, P_5, P_6]$, is employed in the realized system. These are realized with six ultrasonic modules HC-SR04 and have a central angle of 60° , [16], [42]. Positioning sensor locations are predetermined at $(r_{p1}=0.2cm, \theta_{p1}=30^\circ)$, $(r_{p2}=0.2cm, \theta_{p2}=90^\circ)$, $(r_{p3}=0.2cm, \theta_{p3}=150^\circ)$, $(r_{p4}=0.2cm, \theta_{p4}=210^\circ)$, $(r_{p5}=0.2cm, \theta_{p5}=270^\circ)$ and $(r_{p6}=0.2cm, \theta_{p6}=330^\circ)$, respectively. They have equal sensing radii of $r_{p1}=r_{p2}=r_{p3}=r_{p4}=r_{p5}=r_{p6}=(4-0.2)m$ along with equal respective central angles of $\alpha_{p1}=\alpha_{p2}=\alpha_{p3}=\alpha_{p4}=\alpha_{p5}=\alpha_{p6}=60^\circ$. This results in uniform target positioning zones, expressed mathematically by equation 8.

$$Z_{pj} = \frac{1}{2}(3.8)^2 \cdot \frac{\pi}{3}, \text{ where } j \in (1, 2, \dots, 6). \quad (8)$$

By default, cameras remain at their respective home positions and cover Z_{p2} and Z_{p5} . Their orientations are updated to other active zones according to positioning sensor output. This prototype can cover multiple humans present within the system range. In the case of the presence of multiple humans within the field of view of C_1 , the A_1 rotation is prioritized as Z_{p2} , then Z_{p1} and lastly Z_{p3} . Similarly, for the field of view of C_2 , the A_2 rotation is prioritized as Z_{p5} , then Z_{p4} and lastly Z_{p6} .

The developed system is positioned at the center of a $5 \times 5m^2$ laboratory. The setup is shown in Figure 6. System functionality is tested for a simultaneous surveillance of two humans who enter simultaneously the coverage fields of C_1 and C_2 . When an event is detected, human presence is firstly confirmed via the face detector. Later on, depending on the face detector output, alarms are generated and video clips are recorded.

By analyzing and correlating these videos and alarm stamps, it is noted that actions, alarming and video recording, are executed only when an event of interest occurs. False alarms, generated by event detectors, are suppressed by the face detection classifier. Moreover, cameras successfully follow the intended objects as a function of their occurrence in different locations. The lighting block is triggered only when the luminosity of a sensing environment is lower than a predefined threshold. It remains disabled otherwise. It adds to system power consumption efficiency when compared to its counterparts [27], [28], [37].

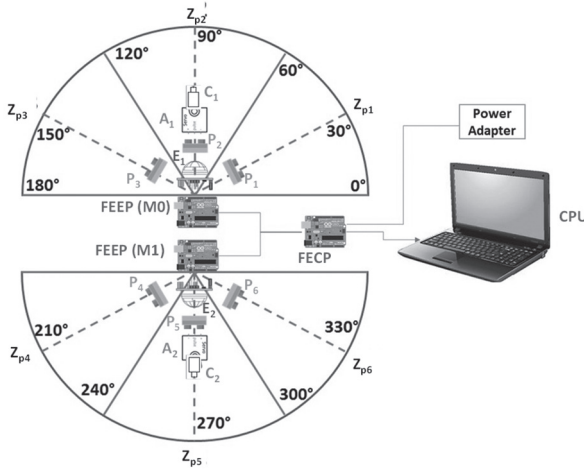


Fig. 5. The coverage principle of the proposed system

System performance is studied in terms of memory utilization. It is accomplished with a three-hour experiment. Throughout this period two persons simultaneously entered the surveillance fields of both front-end modules thirteen times. They did it after random time intervals and at different positioning zones. Each time, they entered the range of the system for a duration of less than two minutes. When an event and the face are detected, a two-second video clip is recorded for the concerned camera. As a result, twenty-six minute videos are recorded, for both C_1 and C_2 , for the three-hour experiment. The process is further clear from results, as shown in Figure 7.

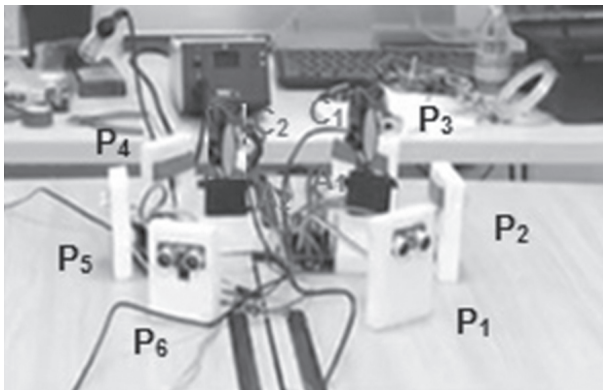


Fig. 6. The proposed system front-end module setup

$$\text{Memory Utilization (Bytes/Sec.)} = \text{Frame size} \times \text{Frame rate} \quad (9)$$

In equation 9, the *Frame Rate* represents the number of frames per second. The frame size can be computed by equation 10.

$$\text{Frame size (Bytes)} = (\text{Camera Resolution} \times \text{Colour Depth})/8 \quad (10)$$

On the right-hand side of equation 10, a constant denominator is employed to calculate the *Frame size* in bytes. By replacing the employed camera resolution and color depth from [49], the *Frame size* of 900KBytes is obtained.

Calculations based on equations 9 and 10 result in a memory utilization of 528MBytes/minute. Therefore, for the proposed system the total video size obtained for each camera is 13.4GB. Therefore, the overall video size obtained for both cameras is 26.8GB.

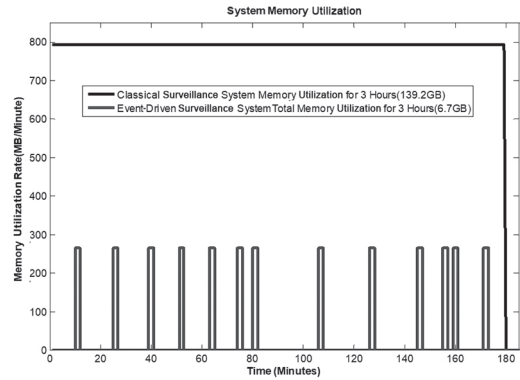


Fig. 7. Surveillance system memory utilization. A black rectangle shows memory utilization for one camera of the classical surveillance system for a duration of 3 hours. Blue pulses show memory utilization for one camera of the proposed event-driven surveillance system for a duration of 3 hours

These results are compared with the counter classical solution. In this case, cameras remain static. Therefore, the surveillance pattern and the field of view of each camera remain fixed. To study the traditional surveillance approach, a system is developed based on the same LifeCam HD-3000 cameras. Each of these cameras can cover a planner angle of 70 degrees. Therefore, 6 cameras are employed for a proper coverage of 360 degrees of the planner surveillance angle. These cameras are connected to the CPU via two USB hubs [48]. All system modules function continuously irrespective of the existence of intended objects in the system range. Moreover, the recording of videos is performed at a rate of 30frames per second. It results in a memory utilization rate of 1584MBytes/minute. The process is clear from Figure 7. The recorded video duration remains one 180 minutes for each camera and at a rate of 15frames per second it delivers into a video size of

278.4GB for each employed camera. Therefore, the overall memory utilization for all cameras is 1670.4GB.

It shows a 20.8 times system memory utilization reduction per camera and an overall 62.3 times system memory utilization reduction achieved by the proposed system when compared to its counterparts [37]. It promises a gain of more than one order degree in terms of power consumption and data processing activity for the proposed system when compared to the traditional one. It is accomplished by efficiently merging the appealing features of event-driven awakening, soft decimation and self-organization techniques.

In order to improve system memory utilization the acquired videos are compacted with the MPEG4 video compressor [50]. This results in a memory utilization rate of 1.4Mbits/second for the proposed solution and of 12.2Mbits/second for the classical one. An appropriate network bandwidth is required for a proper transmission of the surveillance data towards a remote observer. Within this framework, a study is performed in terms of network bandwidth requirements. The key factors which should be taken into account while doing these calculations are the resolution of video recording, the frame rate per second and the employed video compression technique. For real-time transmission a data rate of 1.4Mbits/second is required for the proposed solution. The bandwidth requirement for real-time transmission in the classical case is 12.2Mbits/second. It shows a 9 times bandwidth consumption reduction per second achieved by the proposed solution compared to the traditional one. It is attained by reducing the frame rate and the number of required cameras in the devised system compared to the classical one. Furthermore, in the classical case, the system will continuously consume the network bandwidth at a fixed rate. It results in a bandwidth consumption of 16.1GBytes for the classical case for the overall experiment duration. On the other hand, because of its event-driven and self-organization features, the proposed system does not consume the network bandwidth in a continuous fashion. Bandwidth utilization will be correlated with the intended activity because the proposed system only records and transmits relevant information. The process is clear from Figure 7. Therefore, for overall experiment duration of 3 hours, the proposed system will consume a total bandwidth of 264MBytes. It shows that the devised system achieves an overall bandwidth consumption reduction of 62.3 times with respect to its counterparts.

6. CONCLUSION

An inventive event-driven human surveillance system is proposed in this article. The traditional surveillance systems are constrained because of a prefixed pattern of monitoring. Consequently, false alarms can be triggered and this can result in a reduction of system reliability. Furthermore, all system modules function continuously irrespective of the existence of target

in the range of the system. This causes an augmented processing activity of the system and leads towards an unserviceable utilization of resources and an unwanted consumption of power. With experimental results, it is demonstrated that the proposed system overcomes these downsides up to a certain extent by smartly integrating the event-driven awakening and self-organization features in the proposed solution.

Results have shown that the developed system is capable of monitoring multiple objects. Furthermore, for the case under study, the total 62.3 times reduction of the system memory usage and of the bandwidth consumption is attained by the devised solution in comparison to its counterparts [27], [28], [37]. It tends to achieve a gain of more than one order degree in terms of power consumption of the proposed system when compared to some other ones. It is attained by intelligently integrating event-driven awakening, the soft decimation and the self-organization techniques in the proposed system.

The proposed system power consumption measurement and its comparison with the classical one is in progress. Addition of wireless and mobile vision nodes is also under study. The devised system prototyping and testing with an extended number of front-end modules will be done in the future. A study on system performance in the crowded and real environment is also a future task. The devised system possesses a capability to suppress false alarms. It is attained by intelligently using a robust classifier with front-end sensor arrays. System performance evaluation in terms of false alarm rejection is another future task. An automated alteration of the video recording frame rate as a function of the system environment status is a prospect. It could be realized by integrating machine learning tools in the system. Another prospect is to realize a completely embedded and cost-effective implementation of the devised system.

7. ACKNOWLEDGEMENT

Authors are thankful to anonymous reviewers for their valuable feedback. Authors are also thankful to Eng. S. Ammari for her help during system prototyping and experimentation. The project is funded by the Effat University, Jeddah, KSA.

8. REFERENCES:

- [1] T. Arikuma, Y. Mochizuki, "Intelligent multimedia surveillance system for safer cities", *APSIPA Transactions on Signal and Information Processing*, Vol. 5, 2016, pp. 1-8.
- [2] B. W. Reynolds, "A situational crime prevention approach to cyberstalking victimization: Preventive tactics for Internet users and online place managers", *Crime Prevention and Community Safety*, Vol. 12, No. 2, 2010, pp. 99-118.
- [3] O. P. Popoola, K. Wang, "Video-based abnormal human behavior recognition—A review", *IEEE Transactions on*

- Systems, Man, and Cybernetics, Part C, Vol. 42, No. 6, 2012, pp. 865-878.
- [4] R. Shamosh, H. N. Berlent, U.S. Patent No. 5144661. Washington, DC, 1992, U.S. Patent and Trademark Office.
- [5] Y.Wang, Y. Hu, X. Hu, Y. Zhao, "A Vision-Based Method for Parking Space Surveillance and Parking Lot Management", Proceedings of the International Conference on Image and Graphics, Tianjin, China 13-16 August 2015, pp. 516-528.
- [6] R. Sundar, S. Hebbar, V. Golla, "Implementing intelligent traffic control system for congestion control, ambulance clearance, and stolen vehicle detection", IEEE Sensors Journal, Vol. 15, No. 2, 2015, pp. 1109-1113.
- [7] S. Mukherjee, "Automated Railway Surveillance System", Journal of Applied and Fundamental Sciences, Vol. 1, No. 2, 2015, pp. 268-273.
- [8] Video-based surveillance systems: computer vision and distributed processing, G. A. Jones, N. Paragios, C. S. Regazzoni, (Eds.), Springer Science & Business Media, 2012.
- [9] N. K. Suryadevara, S. C. Mukhopadhyay, "Wireless sensor network based home monitoring system for wellness determination of elderly", IEEE Sensors Journal, Vol. 12, No. 6, 2012, pp. 1965-1972.
- [10] I. Ahmad, "Multi-view video: get ready for next-generation television", IEEE Distributed Systems Online, Vol. 8, No. 3, 2007, p. 6.
- [11] D. Ehret, A. Lau, S. Bittman, W. Lin, T. Shelford, "Automated monitoring of greenhouse crops", Agronomie, Vol. 21, No. 4, pp. 403-414.
- [12] S. Boragno, B. Boghossian, J. Black, D. Makris, S. Velastin, "A DSP-based system for the detection of vehicles parked in prohibited areas", Proceedings of the 2007 IEEE Conference on Advanced Video and Signal Based Surveillance, London, UK, 5-7 September 2007, pp. 260-265.
- [13] A. C. Caputo, "Digital video surveillance and security", Butterworth-Heinemann, 2014.
- [14] A. Taha, H. Zayed, M. E. Khalifa, M El.-S. El-Horbaty, "Human activity recognition for surveillance applications", Proceedings of the 7th International Conference on Information Technology, Amman, Jordan, 12-15 May 2015, pp. 577-586.
- [15] G. Thiel, "Automatic CCTV surveillance-towards the VIRTUAL GUARD", IEEE Aerospace and Electronic Systems Magazine, Vol. 15, No. 7, 2000, pp. 3-9.
- [16] N. Al-Yamani, S. Qaisar, A. Alhazmi, S. Mohammad, A. Subasi, "An event driven surveillance system", Proceedings of the 5th International Conference on Electronic Devices, Systems and Applications, Ras Al Khaimah, United Arab Emirates, 6-8 December 2016, pp. 1-4.
- [17] S. M. Qaisar, R. Yahiaoui, T. Gharbi, "An efficient signal acquisition with an adaptive rate A/D conversion", Proceedings of the 2013 IEEE International Conference on Circuits and Systems, Kuala Lumpur, Malaysia, 18-19 September, pp. 124-129.
- [18] S. M. Qaisar, D. Dallet, S. Benjamin, P. Desprez, R. Yahiaoui, "Power efficient analog to digital conversion for the Li-ion battery voltage monitoring and measurement", Proceedings of the 2013 IEEE International Instrumentation and Measurement Technology Conference, Minneapolis, Minnesota, USA, 6-9 May 2013, pp. 1522-1525.
- [19] S. M. Qaisar, L. Fesquet, M. Renaudin, "Adaptive rate filtering a computationally efficient signal processing approach", Signal Processing, Vol. 94, 2014, pp. 620-630.
- [20] G. Nicolis, I. Prigogine, "Self-organization in non-equilibrium systems", Wiley, New York, 1977.
- [21] B. Berlin, "Ethnobiological classification: Principles of categorization of plants and animals in traditional societies", Princeton University Press, 2014.
- [22] L. Shen, et al., "Large-Scale Far-Infrared Invisibility Cloak Hiding Object from Thermal Detection", Advanced Optical Materials, Vol. 3, No. 12, 2015, pp.1738-1742.
- [23] A. Kuze, H. Suto, M. Nakajima, T. Hamazaki, "Thermal and near infrared sensor for carbon observation Fourier-transform spectrometer on the Greenhouse Gases Observing Satellite for greenhouse gases monitoring", Applied optics, Vol. 48, No. 35, 2009, pp. 6716-6733.
- [24] P. Cheng, F. Zhang, J. Chen, Y. Sun, X. Shen, "A distributed TDMA scheduling algorithm for target tracking in ultrasonic sensor networks", IEEE Transactions on Industrial Electronics, Vol. 60, No. 9, 2013, pp. 3836-3845.
- [25] W. Niem, W. Niehsen, H. Loos, U.S. Patent No. 9,025,033. Washington, DC, 2015, U.S. Patent and Trademark Office.

- [26] J. Liu, S. Sridharan, C. Fookes, T. Wark, "Optimal camera planning under versatile user constraints in multi-camera image processing systems", *IEEE Transactions on Image Processing*, Vol. 23, No. 1, 2014, pp. 171-184.
- [27] T. Soumya, S. M. Thampi, "Day color transfer based night video enhancement for surveillance system", *Proceedings of the 2015 IEEE International Conference on Signal Processing, Informatics, Communication and Energy Systems*, Kozhikode, India, 19-21 February 2015, pp. 1-5.
- [28] S. Zhang, E. Staudt, T. Faltemier, A. K. Roy-Chowdhury, "A camera network tracking (CamNeT) dataset and performance baseline", *Proceedings of the 2015 IEEE Winter Conference on Applications of Computer Vision*, Waikoloa, Hawaii, USA, 5-9 January 2015, pp. 365-372.
- [29] A. Faro, D. Giordano, C. Spampinato, "Adaptive background modeling integrated with luminosity sensors and occlusion processing for reliable vehicle detection", *IEEE Transactions on Intelligent Transportation Systems*, Vol. 12, No. 4, 2011, pp. 1398-1412.
- [30] J.-B. Huang, M.-L. Duan, "2C-bus and the application in microcomputer MCS-51 system", *Journal of Shaanxi University of Science & Technology*, Vol. 1, 2008, p. 23.
- [31] Y. Margalit, D. Margalit, R. Kastershtien, "USB key apparatus for interacting with a USB host via a USB port", U.S. Patent No. US6763399B2. Washington, DC 2004, U.S. Patent and Trademark Office.
- [32] Y.-Q. Wang, "An analysis of the Viola-Jones face detection algorithm", *Image Processing On Line*, Vol. 4, 2014, pp. 128-148.
- [33] S. Agrawal, P. Khatri, "Facial expression detection techniques: based on Viola and Jones algorithm and principal component analysis", *Proceedings of the 5th International Conference on Advanced Computing & Communication Technologies*, Haryana, India, 21-22 February 2015, pp. 108-112.
- [34] H. Sager, W. Hoff, "Pedestrian detection in low resolution videos", *Proceedings of the IEEE Winter Conference on Applications of Computer Vision*, Steamboat Springs, Colorado, USA, 24-26 March 2014, pp. 668-673.
- [35] E. Yusufli, R. Yahiaoui, S. M. Qaisar, T. Gharbi, "An FPGA Based Resources Efficient Solution for the OmniVision Digital VGA Cameras Family," *Proceedings of the HPCC, CSS, ICESS IEEE Conferences*, Paris, France, 20-22 August 2014, pp. 558-561.
- [36] T. Soumya, S. M. Thampi, "Self-organized night video enhancement for surveillance systems", *Signal, Image and Video Processing*, Vol. 11, No. 1, 2017, pp. 57-64.
- [37] E. Ristani, F. Solera, R. Zou, R. Cucchiara, C. Tomasi, "Performance measures and a data set for multi-target, multi-camera tracking", *Proceedings of the European Conference on Computer Vision*, Amsterdam, The Netherlands, 8-16 October 2016, pp. 17-35.
- [38] A. D'Ausilio, "Arduino: A low-cost multipurpose lab equipment", *Behavior research methods*, Vol. 44, No. 2, 2012, pp. 305-313.
- [39] Y. A. Badamasi, "The working principle of an Arduino", *Proceedings of the 11th International Conference on Electronics, Computer and Computation*, Abuja, Nigeria, 29 September - 1 October 2014, pp. 1-4.
- [40] R. C. Brown, "IRIS: Intelligent Roadway Image Segmentation," *Virginia Tech*, <https://vtechworks.lib.vt.edu/handle/10919/49105> (accessed: 2018)
- [41] O. Fratu, S. Halunga, R. Crăciunescu, A. Vulpe, "Instrumental IoT-from environmental monitoring to cosmic ray detection", *Proceedings of the 22nd Telecommunications Forum Telfor*, Belgrade, Serbia, 25-27 November 2014 pp. 17-22.
- [42] Y. Z. Hou, "Assistive guide robot for visually impaired person", *Universiti Teknologi Malaysia*, Faculty of Electrical Engineering, Bachelor thesis, 2013.
- [43] L. Maddalena, A. Petrosino, "A self-organizing approach to background subtraction for visual surveillance applications", *IEEE Transactions on Image Processing*, Vol. 17, No. 7, 2008, pp. 1168-1177.
- [44] C. Rougier, J. Meunier, A. St-Arnaud, J. Rousseau, "Robust video surveillance for fall detection based on human shape deformation", *IEEE Transactions on Circuits and Systems for Video Technology*, Vol. 21, No. 5, 2011, pp. 611-622.

- [45] R. Beaver, et al., "Using degree of confidence to prevent false security system alarms", U.S. Patent Application 14/827715, filed August 17, 2015.
- [46] T. L. Norman, "Integrated Security Systems Design: A Complete Reference for Building Enterprise-wide Digital Security Systems", Butterworth-Heinemann, 2014.
- [47] J. Silverman, M. D. Rostoker, U.S. Patent No. 6370603B1, Washington, DC, 2002 U.S. Patent and Trademark Office.
- [48] T. N. Nguyen, H. T. Nguyen, "Real-time video streaming with multi-camera for a telepresence wheelchair", Proceedings of the 14th International Conference on Control, Automation, Robotics and Vision, Phuket, Thailand, 13-15 November 2016, pp. 1-5.
- [49] "LifeCam HD-3000", Technical Data Sheet, Microsoft Corporation, 2014.
- [50] I. E. H. Richardson, "264 and MPEG-4 video compression: video coding for next-generation multimedia", John Wiley & Sons, 2004.