مجلة العلوم البحثة والتطبيقية
Journal of Pure & Applied Sciences
www.Suj.sebhau.edu.ly ISSN 2521-9200
Received 03/12/2017 Revised 08/02/2018 Published online 30/06/2018

# Secure remote control of offices and stocks based on use of CCTV and image process

*Ahmed B. Elmadani [a] , Abdulkarim I. Abughfa [b]
[a] Department of Computer Science, Faculty of Information Technology, Sebha University, Libya
[b] Department of Electrical and Computer Engineering, School of Applied Sciences and Engineering, Libyan Academy, Libya
*Corresponding Author: ahm.ali@sebhau.edu.ly

**Abstract** Offices and related stocks day to day are in increase to several and they are distributed in deferent places, spotting them become a tedious job. CCTV and internet, and other technologies were used to facilitate monitoring them, but most purposes require that humans detect and respond to events and due to that error are always exist. This paper proposes an algorithm uses secure internet and image process to control government and company branches. Meanwhile it solves mobility and minimizes human intervention in decision making regarding received scene.

**Keywords**: Image process, Hu moment, Mean squire error (MSE), Decision, and Mobility.

# المراقبة الامنة للمكاتب والمخازن عن بعد باستخدام كاميرات المراقبة ومعالجة الصورة

*أحمد بابا المدني علي[1] و عبدالكريم ابراهيم ابوغفة[2]
[1] قسم علوم الحاسوب كلية تقنية المعلومات جامعة سبها، ليبيا
[2] قسم الهندسة الكهربية والحاسوب كلية العلوم التطبيقية والهندسة الأكاديمية الليبية طرابلس، ليبيا
*للمراسلة: ahm.ali@sebhau.edu.ly

**الملخص** المؤسسات ومخازنها يوما بعد يوم تزداد اعدادها ورقعتها الجغرافية مراقبتها بشكل امرا صعبا. لقد استخدمت الكاميرا والانترنيت وتقنيات اخري بغرض تسهيل المراقبة الا ان معظم الاساليب المستخدمة تعتمد على التدخل البشري في تقدير درجة الخطورة واتخاذ القرار اللازم وعادة ما يصاحب القرار المتخذ أخطاء بشرية، وتفاديا لذلك تقدم هذه الورقة نموذجا لخوارزمية تعتمد الانترنيت ومعالجة الصورة لمراقبة المؤسسة او الشركة وفروعها كما تومن الحركة وتقليل الاخطاء البشرية باعتماد الية اتخاذ القرار تستند معالجة الصور السابقة والمتخذة انيا.

**الكلمات المفتاحية:** معالجة الصورة، عزم هوف مان، معيار الخطاء مين إسكويرارور، القرار ، الازاحة.

## Introduction

Many surveillance systems are developed to monitor the desired object or locations [9][10]. The surveillance of home can be performed from anywhere with the help of internet and the system through which we will be able to control the surveillance systems from any place on the earth is needed [6]. Video surveillance is an active area in computer vision which plays an important role in many applications, including industrial automation, robotics, autonomous vehicle navigation and human machine interfaces [7]. Close circuit television (CCTV) is used to monitor and record images of what takes place in specific locations in real time [1]. A surveillance camera can operate a year without a toilet break [8]. The CCTV camera has become a crucial instrument for crime control and has enhanced the safety of publics. CCTV camera also helps in reducing crime as it can provide visual evidences. Furthermore, CCTV camera is used by government to monitor highway traffic, to assess the scene of an accident and other public safety purposes. However, due to improper and ineffective variables [15]. Tracking crime or intrusion guidelines on CCTV camera installation in public area, it is uncertain if there is sufficient coverage by the CCTV camera. Instead of a robust approach, the adequacy of CCTV coverage is often determined based on design experience and trial-and-error. Inadequacy coverage will further limit the efficient combination of control and display equipment, as well as the operators' ability to manage the video surveillance system [16]. Despite increased usage of CCTV, and technological advances of surveillance systems, world-wide more than 100 million surveillance cameras in use say [5], but there is still relatively little understanding of how well it actually works for the many purposes for which it is deployed [4]. Most purposes require that humans detect and respond to events in an appropriate and timely fashion [4]. Because most of CCTV systems use cameras fixed in static position they are able only to monitor a limited area [2]. Even with use of a network of CCTV and high digital technology [13].

The problem occurs from the current system is that the viewing angle of the camera is limited due to stationary position of camera, several cameras are used to cover the whole area [12]. Some research suggests changing of camera direction remotely by use of a Bluetooth mobile application [2]. It is important to identify whether CCTV systems and applications meet stakeholder goals [11], and support human operators effectively, in attaining the goals for which the system is set up [3]. The fundamental part of the CCTV system is a reliable image evaluation by a human observer, whose effectiveness is influenced by many through a CCTV, dependent on human intervention and the success of it depends on the alacrity and the alertness of the personnel involved [14]. Demand from governmental or private business to develop smart surveillance system and to introduce new programs and technologies that can minimize error caused by human [16]. Next paragraph proposes an improvement of a CCTV control system that uses a secure surveillance system based on use of image process to help people in decision making.

**Methodology**

**The System:** The following paragraphs will explain equipment and used codes in the proposed system.

**Used Equipment and Tools:** In our experiment five locations (two offices, two inventories and one parking care) were chosen. In each location two surveillance cameras were fixed to capture scene. Personal computer (PC) and Internet connection used to receive and store images. MathLab image processing tool box used for image process, and written code in VBasic Language used for image features extraction and comparison purpose.

**Data collection:** Entrance and selected inside buildings are places chosen to fix cameras. Then images of chosen places are smoothed, filtered with winner filter. Features of each image are then stored in database for farther use.

**Sequence of the proposed system**

The proposed system follows these steps:

- Image receiving
- Image Feature extraction
- Image comparison
- Decision based on comparisons
- Report Alarm to control Room

As shown in Fig. 1, the proposed system receives an image from camera and process it the compare it with stored in database in order to decide and report the right action.

**Image receiving:** As mentioned in paragraph 2.3, each camera sends images of the location where it was fixed to the PC via internet connection. Received images are smoothed, filtered with winner filter in order to come out with its feature.

**Image Feature extraction:** Hu moment was used to extract image feature. Hu moment is a 2-D continuous function *f*(*x*,*y*), the moment of order (*p*+*q*) that is calculated using equation (1) as following:

$$m_{pq} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^p y^q f(x,y)\,dx\,dy \quad \text{for } p,q = 1,2,3,.. \qquad (1)$$

The moment sequence is defined by $m_{pq}$ . These moments are not invariant to translation, rotation and scaling. The central moments of an image is found using the equation (2).

$$\mu_{pq} = \sum_x \sum_y (x-\bar{x})^p (y-\bar{y})^q f(x,y)$$

p,q = 0,1,2,… (2)

where $\bar{x} = \frac{m_{10}}{m_{00}}$ and $\bar{y} = \frac{m_{01}}{m_{00}}$

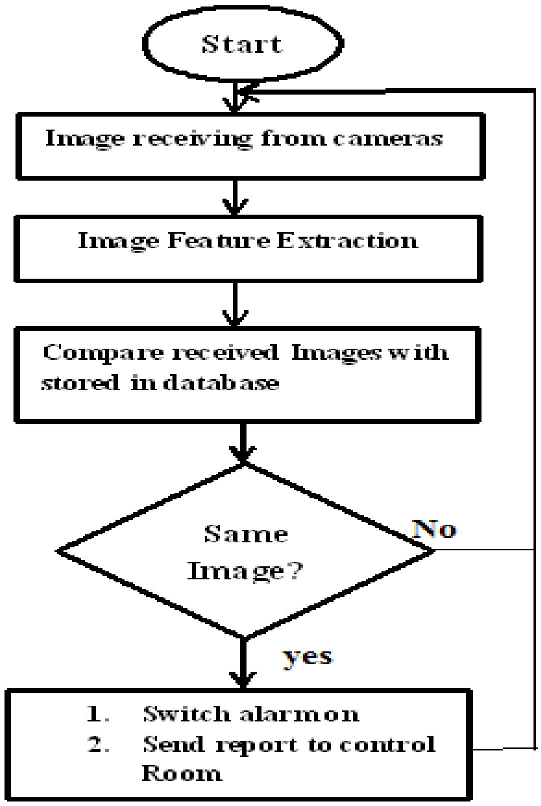


**Fig. 1: Shows sequence of the proposed system.**

The central moments of order up to 3 is found using the equation (3)

$$\mu_{00} = \sum_x \sum_y (x-\bar{x})^0 (y-\bar{y})^0 f(x,y) = \sum_x \sum_y f(x,y) = m_{00}$$

$$\mu_{10} = \sum_x \sum_y (x-\bar{x})^1 (y-\bar{y})^0 f(x,y) = m_{10} - \frac{m_{10}}{m_{00}}(m_{00}) = 0$$

$$\mu_{01} = \sum_x \sum_y (x-\bar{x})^0 (y-\bar{y})^1 f(x,y) = m_{01} - \frac{m_{01}}{m_{00}}(m_{00}) = 0$$

$$\mu_{11} = \sum_x \sum_y (x-\bar{x})^1 (y-\bar{y})^1 f(x,y) = m_{11} - \frac{m_{10}m_{01}}{m_{00}}.$$

$$= m_{11} - \bar{x}m_{01} = m_{11} - \bar{y}m_{10}$$

(3)

The central moments of order up to 3 are found using equations (4)

$$\mu_{20} = \sum_x \sum_y (x-\bar{x})^2 (y-\bar{y})^0 f(x,y) = m_{20} - \bar{x}m_{10}$$

$$\mu_{02} = \sum_x \sum_y (x-\bar{x})^0 (y-\bar{y})^2 f(x,y) = m_{02} - \bar{y}m_{01}.$$

$$\mu_{21} = \sum_x \sum_y (x-\bar{x})^2 (y-\bar{y})^1 f(x,y) = m_{21} - 2\bar{x}m_{11} - \bar{y}m_{20} + 2\bar{x}m_{01}$$

$$\mu_{12} = \sum_x \sum_y (x-\bar{x})^1 (y-\bar{y})^2 f(x,y) = m_{12} - 2\bar{y}m_{11} - \bar{x}m_{02} + 2\bar{y}m_{10}$$

$$\mu_{30} = \sum_x \sum_y (x-\bar{x})^3 (y-\bar{y})^0 f(x,y) = m_{30} - 3\bar{x}m_{20} + 2\bar{x}^2 m_{10}$$

$$\mu_{03} = \sum_x \sum_y (x-\bar{x})^0 (y-\bar{y})^3 f(x,y) = m_{03} - 3\bar{y}m_{02} + 2\bar{y}^2 m_{01}$$

(4)

To normalize the central moments the proposed system uses the following equation:
A seven invariant moments can be derived from the second and third moments using equations (5) and (6).

$$\begin{aligned}
\phi_1 &= \eta_{20} + \eta_{02} \\
\phi_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\
\phi_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\
\phi_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\
\phi_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})\left[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2\right] \\
&\quad + (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})\left[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2\right] \\
\phi_6 &= (\eta_{20} - \eta_{02})\left[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2\right] \\
&\quad + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\
\phi_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})\left[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2\right] \\
&\quad + (3\eta_{12} - \eta_{30})(\eta_{21} + \eta_{03})\left[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2\right]
\end{aligned} \tag{5}$$

$$\eta_{pq} = \frac{\mu_{pq}}{\mu_{00}^{\gamma}} \tag{6}$$

$$\text{where} \quad \gamma = \frac{p+q}{2} + 1 \quad \text{for} \quad p + q = 2, 3, \ldots$$

Hu's seven moment invariants are invariant to image transformations including scaling, translation and rotation. However, this set of moment invariants is not invariant to contrast changes. Hu's seven moment invariants have been widely used in pattern recognition, and their performance has been evaluated under various deformation situation including blurring spatial degradations random noise skew and perspective transformations. As Hu's seven moment invariants take every image pixel into account, the computation cost will be much higher than boundary-based invariants. As stated before, image's spatial resolution decides the total amount of pixels, and to reduce the computation cost, the advantages of traditional moment over other recognition features in shape representation noted during the experiment as following:

- Less computationally demanding and easy to implement.
- Use a single value as the feature, easy for matching.
- Uniqueness.
- Invariant to shape translation, rotation and scaling.
- Less noise-sensitive.

**Assigning Location features to database:** The calculated features of each image location, is to be stored into the database as record. Table 1 shows a sample of a database location feature record.

**Table 1: Location Feature Record**

| No | LN | M1 | M2 | M3 | M4 | M5 | M6 | M7 |
|---|---|---|---|---|---|---|---|---|
| | | | | | | | | |

Where **M1** to **M7** are extracted features and **LN** as location name

Noting that each Mi is calculated using equation (7).

Mi =ABS( (mi1+mi2)/2). (7)

mi1 and mi2 are extracted features for particular location using two cameras within different angles to form moment values of Mi denoted in Table (1) as M1 to M7.

**Feature comparison and action decision**

Mean squire error (MSE) as a measurement of the Euclidian distance between two images is used to equality between the original image and degraded image. It shows if the features of the two images are same or not. The MSE was calculated using equation (8).

$$MSE = \sqrt{\sum_{i=0}^{M}\sum_{m=0}^{N} \left| (\eta_i - \eta_j)^2 \right|} \tag{8}$$

M and N are calculated features of original image and new captured image. $\eta_i$ is the value at moments in the original image where $\eta_j$ values of moments in the corresponding image. The MSE value of two images is calculated by finding the differences between calculated correspondent moments using equation (8). Decision based of resulted MSE values as following:

- Two equal images will give value of MSE zero or close to zero which means no changes in since.
- Two images with value of MSE near or close to 1 means they are different than before so a report should be send to the control room.

**Report Alarm to Control Room:** The MSE value if it scores value equals to 1 or close to 1 an alarm is switched on in the control room and flash lamp indicates the location, that received features are different from the stored in the database. A report is screened for that particular location.

**Testing and Results:** To test the proposed system, two video surveillance cameras were fixed in two different locations. Fig. 2 shows scene of captured images for location L1 and Location L2, the captured scene images are ready for features extraction. Similarly Fig. 3 shows scene images for location L3 and Location L4.


**Fig. 2 Sample of camera's scene for Locations L1 and L2**


**Fig. 3: Sample of camera's scene for locations L3 and L4.**

In Fig. 4 two images for the same location L6, L6A stored image where in L6B shows someone trying to do something in the location L6.


**Fig. 4: L6A and L6B two images for same location (L6)**

Similarly Fig. 5 shows two different images of location L5. L5A as stored in database and L5B as new received image.


**Fig. 5: L5A and L5B two images for same location (L5)**

In Fig.6 shows different images for location L9, L9A represents image as stored in database and image L9B as a changed received image. It requires some action to the control as it will be explained later.


**Fig. 6: L9A and L9B two images for same location (L9)**

The captured image of each location will be smoothed, filtered, given a name, then features are extracted using equations 1 to 7 and stored into database. Equations are used to find Hu moment level one to seven denoted as M1 to M7. Table (2) shows extracted information (M1 to M2) of eleven locations as they stored in the database.

**Table 2: Eleven locations features as stored in database**

| No | LN | M1 | M2 | M3 | M4 | M5 | M6 | M7 |
|---|---|---|---|---|---|---|---|---|
| 1 | L1 | 0.537 | 1.833 | 2.258 | 0.862 | 0.318 | 1.307 | 0.433 |
| 2 | L2 | 1.349 | 3.034 | 0.725 | 0.063 | 0.713 | 0.205 | 0.124 |
| 3 | L3 | 0.671 | 1.207 | 0.717 | 1.630 | 0.488 | 1.034 | 0.726 |
| 4 | L4 | 0.888 | 1.147 | 1.068 | 0.809 | 2.944 | 1.438 | 0.325 |
| 5 | L5 | 0.102 | 0.241 | 0.319 | 0.312 | 0.864 | 0.030 | 0.164 |
| 6 | L6 | 0.863 | 0.077 | 1.214 | 1.113 | 0.006 | 1.532 | 0.769 |
| 7 | L7 | 1.089 | 0.627 | 0.552 | 1.101 | 1.544 | 0.085 | 1.492 |
| 8 | L8 | 0.342 | 0.371 | 3.578 | 2.769 | 0.293 | 1.094 | 1.712 |
| 9 | L9 | 1.489 | 0.742 | 1.409 | 1.417 | 1.370 | 0.226 | 1.109 |
| 10 | L10 | 0.303 | 0.578 | 0.293 | 0.787 | 0.032 | 0.742 | 1.117 |
| 11 | L11 | 0.754 | 1.409 | 1.370 | 2.351 | 0.627 | 1.062 | 0.371 |

The stored information is used for comparison purpose. Times to time stored features are checked to see if there is any changing. The system is configured to send several images then features are extracted from received images. As an example the received image shown in Fig. 3, it is for location L6, one is as stored in database and second deferent, features are extracted and compared. If received features are not the same as stored in database that means scene is changed. In Table (3), the extracted features as seen, are different from the stored in database. A written computer code is used compare features and desired the right action. Equation (8) is used to compare features. Results of comparison will be equal to zero or close to zero if two images are same and no changes done, otherwise close to one or equal to one which denote that changes done to the previous image. By applying equation (8) using received information in L6 it is found that MSE equals to 2.317 and similarly by applying equation (8) to the new information received form location L9 a MSE was found equals to 1.378. So because of the changes noted in locations L6 and L9, a report has to be sent to the control room.

**Table 3: Extracted features of L6 after change**

| No | LN | M1 | M2 | M3 | M4 | M5 | M6 | M7 |
|---|---|---|---|---|---|---|---|---|
| 1 | L9 | 1.491 | 1.872 | 1.710 | 1.937 | 1.770 | 0.451 | 1.333 |

**Table 4: Extracted features of L9 after change**

| No | LN | M1 | M2 | M3 | M4 | M5 | M6 | M7 |
|---|---|---|---|---|---|---|---|---|
| 1 | L6 | 0.541 | 1.893 | 2.458 | 0.886 | 0.538 | 1.707 | 0.526 |

**Result Analysis:** As it is described, the proposed system needs not highly user attention as most of the traditional surveillance system, because it provides an automated remote control without any human intervention. This action is performed in a periodical time. Images are different from stored ones if and only if the values of MSE shown in the column titled by AP and their percentage titled by AP%, are close to one. The features comparisons are performed between pre-stored images in database with features of the received image**.** Computer code were written and used to check for any differences between features of the two images. Traditional surveillance systems based on human intervene and the reliable image evaluation is performed by a human observer, where in the proposed system comparisons of features are performed automatically without any human errors, therefore errorless. Table 5, shows results of testing five locations selected randomly. The system shows number of proper actions performed (AP) and their percentage (PA%) as a result of comparing the stored image and received image. Images are not different as shown in the column titled by AP and it is percentage titled by AP% if thevalues of MSE are close to one so action will be performed. In Table 5, the proposed system shows some non-performed action (NA) and their percentage (NA%) due to the values of MSE which are very close to zero, which means the two images are very similar.

**Table 5: Results for chosen locations randomly**

| No. | LN | TI | AP | AP% | NA | NA% |
|---|---|---|---|---|---|---|
| 1 | L2 | 30 | 29 | 97% | 1 | 3% |
| 2 | L4 | 45 | 44 | 98% | 1 | 2% |
| 3 | L7 | 60 | 58 | 97% | 2 | 3% |
| 4 | L10 | 12 | 12 | 100% | 0 | 0% |
| 5 | L11 | 73 | 70 | 97% | 2 | 4% |

Where LN a location number, TI total received images, AP action performed, AP% percentage of performed action, NA non-performed action, and NA% percentage of non-performed action

**Related Work:** This work is similar to the research given by [2]. The mentioned work uses employee to monitor and analysis received location's images. It is also similar to research given by [7] that sent images through the internet both system are controlled by human being where in the proposed system instead the received images monitored by a computer itself then uses a written code to analysis them to avoid any human mistake.

**Conclusion:** An investigation of surveillance control systems and their drawback were shown in this paper. A system which can operate automatically was proposed. The proposed system represents an avoidance of errors caused by human being automatically.

**References**

[1]- Aus tralian Institute of Criminology; AICrime matters No.18: Closed circuit television (CCTV) as a crime prevention measure, Canberra, AIC. 2004. www.aic.gov.au/publications/crm/crm018.html.

[2]- Patel A, Hema Soni H, Patil S. and Gharge S. (2016).RF Controlled Digital Wireless Camera. International Journal on Recent and Innovation Trends in Computing and Communication Volume: 4 Issue: 4. pp 226-229. ISSN: 2321-8169

[3]- Keval H. (2016). CCTV Control Room Collaboration and Communication: Does it Work. www.cs.ucl.ac.uk

[4]- Keval H.U and Sasse M.A. (2006).Man or a Gorilla? Performance Issues with CCTV Technology in Security Control Rooms. www.cs.ucl.ac.uk

[5]- The Video Surveillance Industry – Market size. (2016) www.cameramanager.com.

[6]- Home surveillance system. (2016) proceedings of International Conference Advances in Computing, Communication, & Automation (ICACCA), Dehradun, India, DOI: 10.1109/ICACCA.2016.7578890.

[7]- Venkatesan R., Dinesh P. A.R. and A. Balaji A. G.(2016). Unsupervised Learning Based Video Surveillance System Established Networked Cameras.

[8]- Adrienne I. (2001).Can Surveillance Cameras Be Successful In Preventing Crime And Controlling Anti-social Behaviours?. The Character, Impact and Prevention of Crime in Regional Australia Conference convened by the Australian Institute of Criminology and held in Townsville. Institute of Criminology and held in Townsville.

[9]- UddUddin M. J. Ullah A. and Hoq M. A (2014). Automatic Wireless Remote Control Car With Surveillance Camera: A

successful research project executed by Department of Electrical & Electronic Engineering (EEE), IIUC Paperback

[10]- Anthony C. Caputo. (2014). Digital Video Surveillance and Security, Published by Elsevier. Inc. Second Edition. ISBN-13: 978-0124200425
ISBN-10: 0124200427

[11] DauDaughert.M.J. (2013). The Devil Inside the Beltway: The Shocking Expose of the US Government's Surveillance and Overreach Into Cybersecurity, Medicine and Small Business Softcover Edition. IISBN-13: 978-0985742225 IISBN-10: 0985742224

[12]- *KruKruegle H and Abram* F. (2016). CCTV Surveillance (Second Edition) Analog and Digital Video Practices and Technology. Elsevier Inc. ISBN: 978-0-7506-7768-4

[13]- Damjanovski V. (2014). CCTV Networking and Digital Technology. Elsevier Inc ISBN: 978-0-12-404557

[14]- Das tidar J. G. and Biswas R. (2015). Tracking Human Intrusion through a CCTV Proceedings of Computational Intelligence and Communication Networks (CICN), International Conference. India. ISSN: 2472-7555-6.
DOI: 10.1109/CCST.2005.1594815.

[15]- Voorthuijsen G. V. , Hoof H. V, K. Roubik K, Bernas M and Pata P. (2005). CCTV effectiveness study. Proceedings of Carnahan Conference Security Technology (CCST) Spain.

[16]- Asirvadam V. S, Fen C. P, Saad N, Sebastian P, Altahir A. A, Ali A and Hamid N. H.(2014). Risk map for video surveillance CCTV. Proceedings International Conference Intelligent and Advanced Systems (ICIAS). Kuala Lumpur, Malaysia.
DOI: 10.1109/ICIAS.2014.6869527.