An intelligent real-time occupancy monitoring system with enhanced encryption and privacy
Ahmad, Jawad; Larijani, Hadi; Emmanuel, R; Mannion, Mike; Javed, Abbas; Ahmadinia, Ali

Published in:
2018 IEEE 17th International Conference on Cognitive Informatics & Cognitive Computing (ICCI*CC)

DOI:
10.1109/ICCI-CC.2018.8482047

Publication date:
2018

Document Version
Author accepted manuscript

Link to publication in ResearchOnline

Citation for published version (Harvard):

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

Take down policy
If you believe that this document breaches copyright please view our takedown policy at https://edshare.gcu.ac.uk/id/eprint/5179 for details of how to contact us.
An Intelligent Real-Time Occupancy Monitoring System With Enhanced Encryption and Privacy

Jawad Ahmad*, Hadi Larijani *, Rohinton Emmanuel *, Mike Mannion *, Abbas Javed †, and Ali Ahmadinia ‡
* School of Engineering and Built Environment, Glasgow Caledonian University, Glasgow, United Kingdom.
† Department of Electrical Engineering, COMSATS Institute of Information Technology, Lahore, Pakistan.
‡ Department of Computer Science, California State University San Marcos, United States.
jawad.ahmad@gcu.ac.uk, H.larijani@gcu.ac.uk, rohinton.emmanuel@gcu.ac.uk,
m.a.g.mannion@gcu.ac.uk, abbasjaved@ciitlahore.edu.pk, and aahmadinia@csusm.edu

Abstract—The number of people entering or leaving a building is an essential piece of information that has a lot of practical applications in intelligent building, queue management, and customer service. Vision-based technologies are widely installed in real-time occupancy monitoring systems due to accuracy and reliability. However, monitoring occupancy through unprotected video may disclose privacy of innocent people. Therefore, protecting confidentiality and accurately counting the number of people in real-time scenarios is a severe challenge. Encrypting such videos is one of the promising solutions for maintaining privacy. In this paper, we propose a real-time occupancy monitoring system with Region of Interest (ROI) based light-weight video encryption. People movement is detected through a widely used background model, i.e., Gaussian Mixture Model (GMM) and Kalman filter. Instead of encrypting the whole frame including background, the main idea is to encrypt people present in video via Tangent Delay Ellipse Reflecting Cavity Map System (TD-ERCS). Compared to existing schemes which are mainly based on complete encryption, the proposed method provides partial encryption though cryptographically secure and low-cost computation. The proposed scheme is tested with several different parameters such as correlation, entropy, contrast, energy, Number of Pixel Change Rate (NPCR) NPCR, Unified Average Change Intensity (UACI) and key space. Results from all security parameters have highlighted sufficient security of the proposed scheme.

Index Terms—Occupancy, video processing, image encryption, chaos, TD-ERCS map

I. INTRODUCTION

MODERN video surveillance monitoring systems can provide an image processing based solution for occupancy system, smart buildings management, retail traffic analysis, and security applications. Video-based occupancy counting systems have several advantages over other traditional methods [1], [2]. Compared to other traditional technologies, camera hardware is low-cost, and the output of video-based occupancy is more precise and considered to be more reliable concerning accuracy. As a result, camera-based people counting systems have gained superiority over other traditional occupancy systems [1]. However, camera-based occupancy systems do not conceal the individual identity and hence causes serious privacy concerns [3], [4]. Without protecting or encrypting any sensitive data, images collected from video-surveillance systems might break country laws on privacy. Due to privacy protection requirements, many researchers are attracted towards designing efficient and secure image and video encryption schemes [3]–[6]. Traditional encryption algorithms such as Advanced Encryption Standard (AES), Rivest Shamir Adleman (RSA), International Data Encryption Algorithm (IDEA) and Data Encryption Standard (DES) are designed for text encryption and cannot be used in video/image encryption due to following reasons:

1. Encrypting video through traditional methods such as AES requires high computing power.
2. Video data are large in volume, and hence traditional scheme are not suited in real-time applications.
3. Traditional encryption schemes are slow and hence cause real-time latency.

Homomorphic encryption is considered an essential tool for processing of sensitive data [7]. Several studies have been carried out on encrypted signal and image processing such as image, encrypted image watermarking, enhancement, encrypted face recognition, and many other image encryption techniques have seen enormous growth in research during the last decade. Due to high data volumes, encrypting video data in real-time is a highly challenging task. Furthermore, the protection of text data is not difficult as compared to the video because video data protection involves the non-disclosure of video content itself which makes video security an essential research topic. Video security is defined by Stutz et al. [8] as protecting video data such that an attacker is not capable of recovering original video content information from the encrypted video with higher quality than accepted in the real-time application. It is evident from this context that other than standard image encryption techniques there are several methodologies which can be adopted to achieve video security because it’s neither efficient nor practical to use conventional cryptosystems to encrypt the full-length video bitstream. Hence, to satisfy real-world video encryption requirements, partial encryption also known as selective encryption is a popular choice of researchers [9]. In selective encryption, only a small portion or subset of data is encrypted [9]. The main aim of such schemes is to encrypt only small amounts of data while still preserving security. Thus selective encryption is computationally efficient and well-suited in real-time applications.
There are several studies available in the literature which primarily focused on the privacy of surveillance videos. The primary objectives were to hide the sensitive data from eavesdroppers. In [10], Newton et al., proposed a novel privacy-enabled method, known as K-name which de-identify facial features and the face of the original person cannot be recognised. Results of Newton et al. scheme were verified in real-time [10]. However, several attacks on the system are possible which have also been reported in [10]. The same methodology can be found in [4], [11]. The primary goal of these studies was invalidating facial recognition and hence protecting sensitive data from an unauthorised person. Only an authorised person can recover the features and can see the original face images in surveillance videos. ROI-based privacy methods blur not only facial characteristics but also blur entire moving objects. In literature, some studies can also be found which blur whole moving objects [6]. Defaux et al., scheme protects the privacy of ROI through either transform-domain or codestream-domain. In transform domain, the sign of some coefficients was randomly reversed. In the second technique, codestream bits were randomly scrambled. Both transform-domain and codestream-domain scrambling were tested only on MPEG-4. Noise or randomness introduced in ROI can be adjusted in Defaux et al., scheme. An unauthorised person cannot see the original ROI objects. In [5], Martin et al. presented a novel algorithm for some particular shapes using Secure Shape and Texture SPIHT (SecST-SPIHT). Less than 5% of codes were encrypted instead of “whole content” encryption in Martin et al., scheme.

The main contributions of this work are: 1. Development of a novel ROI-based light-weight image encryption along with people counting scheme, 2. Critical evaluation in real-world environment, and critical analysis of results. The rest of paper is organised as follows: TD-ERCS chaotic map and GMM are discussed in Section II. Section III explained the proposed scheme in every detail. Results and discussions are outlined in Section IV. Conclusions and future work are drawn in Section V.

II. BACKGROUND AND RELATED WORK

In last two decades, the security researchers have reported a close relationship between chaotic maps and cryptography. Encryption based on chaotic or chaotic maps (for example logistic and tent maps), share similar characteristics. Both are sensitive to initial conditions and difficult to predict its output. Due to highly unpredictable nature and random-like behaviour, cryptographers have designed encryption scheme based on chaotic maps. Xiang et al., proposed a light-weight block cryptographic scheme mainly based on Logistic map [12]. The plaintext blocks were randomly shuffled using binary sequences obtained from the chaotic map. Although the Xiang et al., scheme was light-weight and fast, however, due to the low key space, the proposed scheme was proved to be insecure as outlined in [13]. Wang et al., critically analyses Logistic map based scheme and reported that low key space scheme is insecure against many attacks. A new encryption strategy based on circular inter-intra bit level permutation method was proposed by Diaconu [14]. Some statistical and differential analyses including entropy tests were carried out, and Diaconu proved his novel proposal [14]. But recently Fan et al., reported plaintext attack on Diaconu scheme [15]. Security of original scheme was improved via changing two basic steps permutation and diffusion. The modified version of Diaconu scheme was resistant against plaintext attack [15].

As per web of science record, until 2014, more than two thousands papers were published on chaos-based cryptography. Unfortunately, many of these cryptographic algorithms are either insecure or impractical due to low key space and/or cryptographically insecure as highlighted in [12], [13], [15]–[17]. TD-ERCS map also possesses other properties such as zero correlation in total field, positive Lyapunov exponent and equiprobable distribution [18]. In our proposed scheme, we also selected TD-ERCS map due to large key space and aforesaid properties. TD-ERCS can be written as [19]:

\[
\begin{align*}
    x_n &= \frac{-2k_{n-1}y_{n-1} + x_{n-1}(\mu^2 - k_{n-1}^2)}{\mu^2 + k_{n-1}^2} \\
    y_n &= k_{n-1}(x_n - x_{n-1}) + y_{n-1}, \quad n = 1, 2, 3, \ldots \tag{1}
\end{align*}
\]

where

\[
\begin{align*}
    k_n &= \frac{2k_{n-m} - k_n + k_{n-m}(k_{n-m}^2 - 1)}{1 + 2k_{n-m}k_n - (k_{n-m}^2 - 1)} \tag{2} \\
    k_{n-m} &= \begin{cases} 
        \frac{-x_{n-m}}{y_{n-m}}\mu^2 & n < m \\
        \frac{-y_{n-m}}{x_{n-m}}\mu^2 & n \geq m 
    \end{cases} \tag{3} \\
    y_0 &= \mu\sqrt{1 - x_0^2} \tag{4} \\
    k_0' &= \frac{x_0}{y_0}\mu^2 \tag{5} \\
    k_0 &= \frac{\tan^{-1} k_0' - \mu}{1 - k_0'^2} \tag{6} \\
    \begin{cases} 
        \mu &\in (0, 1) \\
        x_0 &\in [-1, 1] \\
        \alpha &\in (0, \pi) \\
        m &\in 2, 3, 4, 5, \ldots 
    \end{cases} \tag{7}
\end{align*}
\]

Where \( \mu, x_0, \alpha \) and \( m \) are called initial conditions or seed parameters of TD-ERCS map. SDIC property of TD-ERCS map is shown in Fig. 1. From Fig. 1 (a), it is clear that map produces random value after 20th iteration. One can also see from Fig. 1(b) that output also changes if \( \mu \) changes slightly. Similar results can be proved for other parameters. Thus \( x_0, \mu, \alpha \) and \( m \) can be used as initial secret key parameters in real-time video encryption.

Figure 1 highlights SDIC property for the TD-ERCS map. It can be seen from Fig. 1(a) that map values are totally different after 20th iteration. From Fig. 1(b), one can also notice that not only \( x_n \) values differ for different values of \( x_0 \), but slight different values of \( \mu \) also produced different output after some iterations. In subsequent parts of the paper, we will use \( x_0 \) and \( \mu \) as a key parameter.
In order to detect any moving object out of video sequence and then encrypting followed by counting, one must use background modelling. For moving object detection, various background modelling methods can be found in the literature. The background can be detected via one of the well-known method GMM. GMM uses multiple Gaussian functions for the adoption of environmental changes. Mathematically, GMM can be defined as:

$$P(x_t) = \sum_{j=1}^{K} w_{j,t} \eta(x_t; \phi_{j,t}, \sum_{j,t}),$$

(8)

In Eq. 8 $x_t$ is incoming pixel value at a time $t$, $w_{j,t}$ is weight of $j$th distribution at time $t$ and $\eta(x_t; \phi_{j,t}, \sum_{j,t})$ is probability Gaussian distribution function. Based on value of $w/\sigma$, $K$ distribution are sorted and background model can be written as follows:

$$B = \arg \min_b \left( \sum_{j=1}^{b} w_k > T \right),$$

(9)

Each Gaussian which is greater than the threshold $T$ is classified as background. More details about GMM can be found in [3].

Kalman filter is extensively used in the area of control systems, dynamic systems and position estimation. In the proposed method, we used Kalman filter followed by munkres for finding an optimal estimation of person position in current frame and tracking, respectively. Basically, Kalman filter is a set of mathematical equations which provides an efficient position estimation even in presence of Additive White Gaussian Noise (AWGN). A discrete-time Kalman filter for state transition $\hat{X}$ at time $k$ can be written as:

$$\hat{X}_k = A\hat{X}_{(k-1)} + Bu_k + W_k,$$

(10)

where $A$ is $n \times 1$ system state transition vector, $B$ is called control parameter which relates $u_k$ with state $\hat{X}_k$ and $W_k$ is $m \times 1$ unknown process noise. Mathematically, measurement $Z$ in terms of state $\hat{X}_k$ can be defined as:

$$Z_k = H_k\hat{X}_k + V_k,$$

(11)

where $Z$ is measurement, $H$ is the relationship between the measurement $Z$ and the state $\hat{X}_k$ and $V_k$ is measurement noise. Kalman filter basically predicts the next state via the current set of observations and then update the current set of predicted measurements [20]. These basic steps of Kalman filter are also highlighted in Fig. 2 [20]. Kalman filter and iterative Hungarian algorithm are used in this work for tracking.

### III. THE PROPOSED COUNTING AND VIDEO ENCRYPTED SCHEME

In this Section, we present our model which follows four steps i.e., (i) detection, (ii) ROI-based encryption, (iii) tracking and finally (iv) counting number of persons from the video sequences. The goal of our proposed scheme is not to count persons only but to encrypt video sequences in the encrypted domain in real-time scenarios. In order to overcome unnecessary overheads, real-time processing requires light-weight encryption schemes. Block diagram of the proposed method is shown in Fig. 3. Steps are shown in Fig. 3 and explained as:

**Step 1:** Acquired video frames $F$ in real-time using a single overhead camera. To simplify the algorithm, only grayscale image frames are acquired from the camera or convert RGB...
image frames to grayscale.

**Step 2:** Getting a fixed background $B$ for foreground detector is difficult due to the illumination changes, addition or removal of objects in the scene and/or camera movement. In this work, we used a robust method known as GMM as a foreground detector. In this work, we set the number of Gaussian (K) and threshold ($\tau$) as 3 and 0.7, respectively.

**Step 3:** Shadows are generally misclassified as foreground pixels $P$. Remove shadows from the current frames $F$ by converting the RGB image to YCbCr format $Y$. Apply morphological closing and opening operations on $Y$. Mathematically, closing $PC$ and opening operation $PO$ can be defined as:

\begin{equation}
PC = \text{Erode}(\text{Dilate}(P), Kr),
\end{equation}

\begin{equation}
PO = \text{Dilate}(\text{Erode}(PC), Kr),
\end{equation}

here $Kr$ is the kernel or structuring element.

**Step 4:** Based on a threshold $\tau$, current image $F$ is subtracted from background image $B$ to get the foreground moving objects $R$ (ROI) such as person(s). Pixels which are greater than $\tau$ are viewed as foreground pixels.

**Step 5:** Select initial key parameters for TD-ERCS map $\mu, x_0, \alpha$ and $m$.

**Step 6:** Iterate TD-ERCS map $h \times w$ times, where $h \times w$ is the size of ROI bounding box.

**Step 7:** Detected person(s), are encrypted as:

i. In order to overcome transient effects in TD-ERCS map, select last $h + w$ elements in $X$. To get permuted values $P_{ROI}$, randomly permute rows and columns of ROI (detected person(s)). This step of encryption is also known as confusion.

ii. Diffuse values obtained in $P_{ROI}$ via XOR operation. Bitwise XOR $P_{ROI}$ with values in $Y$ and get encrypted ROI.

**Step 8:** People are tracked using Kalman filter prediction model using $X_k$ as outlined in the previous section. Next, the Hungarian cost matrix $\psi$ is calculated with the cost of assignment between propagated track $T$ and every detection $D$.

**Step 9:** In last, count the number of the people person(s) in encrypted domain.

**IV. EXPERIMENTAL RESULTS AND SECURITY ANALYSIS**

Figure 4 highlights the original image and its encrypted version respectively. One can see that the proposed scheme conceal all important information (see parts (c), (e) and (g) in Fig. 4) Only authorised person can access or decrypt the original contents of the video who have the correct TD-ERCS key. Decryption is the reverse process of encryption. Although, the content is encrypted and one cannot see the original information. However, still, the proposed algorithm must be tested and evaluated via some important parameters discussed in Ref [21]. Hence the effectiveness of the proposed scheme is also tested using important security parameters such as correlation contrast, correlation, energy, key space and key sensitivity which are thoroughly defined in our previous work [22], [23]. Due to space limits, parameters mentioned above are not defined in this paper. How these parameters can be used to judge the security of an encrypted scheme can be found in [22], [23].

The low value of correlation coefficient reflects higher security of an image encryption scheme. From Tables I and II, one can see low values of correlation coefficients as compared to the original frame. In Fig. 4 (c), (e) and (g) only encrypted part is cropped from the encrypted frame and correlation coefficient is also calculated for the cropped image in all direction. Correlation coefficient values from Tables I and II highlights that the encrypted part (cropped image) has lowest correlation in all direction i.e., Horizontal ($H_{CC}$), Vertical ($V_{CC}$), Diagonal ($D_{CC}$). Higher the values of entropy better are the encryption algorithm. One can see the entropy values from Tables I and II and can conclude that the proposed scheme offers higher entropy in the ROI. Higher contrast indicates that text is not homogeneous. Results obtained in these tables show a higher contrast for cropped encrypted images. Energy in image processing is a measure of information. Higher the energy is, higher the information content in the image. In a good cryptosystem, an encrypted information contains less information. In the proposed scheme, the energy of the cropped image is least when compared to other images. NPCR and UACI also reveal variance rate of pixels in encrypted images with slight key changes. NPCR and UACI values in Tables I and II are shown which indicates higher encryption quality and security. As a result, the proposed scheme can be considered sensitive to initial key parameters. An intruder cannot get the original information until he has the exact original key. Key space should be large enough to resist brute force attack. Key space of the proposed scheme is approximately $2^{100}$ which is sufficient for resisting such attacks which can be carried out in low key spaces.
Fig. 4: ROI based Image encryption and counting number of person.
In this paper, we present a novel technique for detecting, tracking, encrypting and then counting the number of people passing a virtual line. The proposed method integrates counting with Region of Interest (ROI) based encryption. Only target moving objects are encrypted in the video through a novel chaotic map known as TD-ERCS. The output of TD-ERCS map is also shown in this work which highlights that the map is very sensitive to initial key parameters and thus an attacker cannot guess the original key. The proposed scheme is tested in real-time in an office environment, and security assessment of the proposed scheme is done using various methods to prove its robustness. The map is very sensitive to initial key parameters and thus an attacker cannot guess the original key. The proposed scheme is also shown in this work which highlights that the map is very sensitive to initial key parameters and thus an attacker cannot guess the original key. The proposed scheme is tested in real-time in an office environment, and security assessment of the proposed scheme is done using various methods to prove its robustness.

### Table I: Security assessment of the proposed scheme: One person detected in frame

<table>
<thead>
<tr>
<th>Security Parameter</th>
<th>Original Frame</th>
<th>Encrypted Frame</th>
<th>Cropped Frame</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_{CC}$</td>
<td>0.8962</td>
<td>0.5315</td>
<td>-0.0362</td>
</tr>
<tr>
<td>$V_{CC}$</td>
<td>0.9180</td>
<td>0.5805</td>
<td>0.1453</td>
</tr>
<tr>
<td>$D_{CC}$</td>
<td>0.8231</td>
<td>0.4626</td>
<td>-0.0175</td>
</tr>
<tr>
<td>Entropy</td>
<td>6.6776</td>
<td>6.9647</td>
<td>7.5987</td>
</tr>
<tr>
<td>Contrast</td>
<td>0.4274</td>
<td>2.8591</td>
<td>14.5487</td>
</tr>
<tr>
<td>Energy</td>
<td>0.9011</td>
<td>0.1992</td>
<td>0.0252</td>
</tr>
<tr>
<td>NPCR</td>
<td>NA</td>
<td>67.0225</td>
<td>96.9920</td>
</tr>
<tr>
<td>UACI</td>
<td>NA</td>
<td>8.9078</td>
<td>40.9215</td>
</tr>
</tbody>
</table>

### Table II: Security assessment of the proposed scheme: two persons detected frame

<table>
<thead>
<tr>
<th>Security Parameter</th>
<th>Original Frame</th>
<th>Encrypted Frame</th>
<th>Cropped Frame</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_{CC}$</td>
<td>0.9169</td>
<td>0.2832</td>
<td>0.0350</td>
</tr>
<tr>
<td>$V_{CC}$</td>
<td>0.9352</td>
<td>0.4422</td>
<td>0.1023</td>
</tr>
<tr>
<td>$D_{CC}$</td>
<td>0.8231</td>
<td>0.2202</td>
<td>0.0240</td>
</tr>
<tr>
<td>Entropy</td>
<td>6.738</td>
<td>7.2150</td>
<td>7.6349</td>
</tr>
<tr>
<td>Contrast</td>
<td>0.4833</td>
<td>5.1095</td>
<td>14.6883</td>
</tr>
<tr>
<td>Energy</td>
<td>0.2658</td>
<td>0.1566</td>
<td>0.0252</td>
</tr>
<tr>
<td>NPCR</td>
<td>NA</td>
<td>64.4734</td>
<td>98.7631</td>
</tr>
<tr>
<td>UACI</td>
<td>NA</td>
<td>12.6859</td>
<td>35.7946</td>
</tr>
</tbody>
</table>

### V. Conclusion

In this paper, we present a novel technique for detecting, tracking, encrypting and then counting the number of people passing a virtual line. The proposed method integrates counting with Region of Interest (ROI) based encryption. Only target moving objects are encrypted in the video through a novel chaotic map known as TD-ERCS. The output of TD-ERCS map is also shown in this work which highlights that the map is very sensitive to initial key parameters and thus an attacker cannot guess the original key. The proposed scheme is tested in real-time in an office environment, and security scheme is verified via some parameters such as correlation, entropy, contrast, energy, number of pixel change rate and unified average change intensity. Results from all parameters mentioned above proved sufficient security of the proposed scheme. The limitation of the proposed scheme is that the speed of encryption decreases with increasing the number of people in the frame. In future, the efficiency of the proposed light-weight system will be tested with the different resolution of images. In the extended version of the paper, we will apply different real-time attack scenarios such as known plaintext attack and chosen plaintext attack etc., on the proposed scheme.

### References