Secure Surveillance Framework for IoT Systems Using Probabilistic Image Encryption

Khan Muhammad, *Student Member, IEEE*, Rafik Hamza, Jamil Ahmad, *Student Member, IEEE*, Jaime Lloret, *Senior Member, IEEE*, Haoxiang Wang, and Sung Wook Baik, *Member, IEEE*

***Abstract*—This paper proposes a secure surveillance framework for Internet of things (IoT) systems by intelligent integration of video summarization and image encryption. First, an efficient video summarization method is used to extract the informative frames using the processing capabilities of visual sensors. When an event is detected from keyframes, an alert is sent to the concerned authority autonomously. As the final decision about an event mainly depends on the extracted keyframes, their modification during transmission by attackers can result in severe losses. To tackle this issue, we propose a fast probabilistic and lightweight algorithm for the encryption of keyframes prior to transmission, considering the memory and processing requirements of constrained devices that increase its suitability for IoT systems. Our experimental results verify the effectiveness of the proposed method in terms of robustness, execution time, and security compared to other image encryption algorithms. Furthermore, our framework can reduce the bandwidth, storage, transmission cost, and the time required for analysts to browse large volumes of surveillance data and make decisions about abnormal events, such as suspicious activity detection and fire detection in surveillance applications.**

***Index Terms*—Industrial Internet of things (IoT), information security, lightweight image encryption, surveillance networks, video summarization.**

I. Introduction

THE recent development in the processing capabilities of smart devices has resulted in intelligent Internet of things (IoT) environments, enabling the connecting nodes to collect, perceive, and analyze necessary data from their surroundings and react accordingly. Wireless multimedia surveillance networks (WMSNs) are part of this IoT-assisted environment, which consists of visual sensors that observe the surrounding environment from multiple overlapping views by continuously capturing images, thereby producing a large amount of visual data with significant redundancy [1]–[3]. It is widely agreed in the research community of surveillance networks that the collected visual data should be processed and only the informative data should be recorded for future usage, such as abnormal event detection, case management, data analysis, and video abstraction. The reason is that sending all the imaging data through the communication lines without processing is impractical because of energy and bandwidth constraints. In addition, it is comparatively difficult and time-consuming for an analyst to efficiently extract actionable intelligence from the sheer volume of surveillance data [4].

Therefore, it is necessary to exploit a mechanism that can collect semantically important visual data autonomously by utilizing the processing and transmission capabilities of modern smart visual sensors. Such a mechanism can make it possible to intelligently select the appropriate view from multiview surveillance data captured by multiple sensors connected via IoT infrastructure. It can facilitate the processing of the collected data in real time so as to send only relevant data to the central storage for future use. Furthermore, it enables surveillance specialists to make timely decisions by analyzing only the representative frames, grasping the pertinent contents of the original lengthy sequence of visual data. Some typical surveillance scenarios highlighting events of interest to us in industrial environments are shown in Fig. 1.

The literature review indicates that WMSN-based monitoring systems have two main requirements: first, robustness; and second, efficient resource utilization [5]. The robustness of the real-time surveillance system is often compromised due to failure of visual sensors caused by human intrusion, technical malfunction, or natural catastrophes. This can be avoided by using a multiview camera WMSN. However, the multiview camera WMSN encounters the problem of full or partial coverage overlaps, producing a large volume of redundant data [6]. This results in unnecessary resource utilization of the network in the processing and transmission of such huge data. Further, the visual data in the WMSN are transmitted wirelessly to a visual processing hub (VPH) and base station (BS). This communication is vulnerable to several security issues. It is, therefore, important to send the imaging data securely to the BS with some security

Manuscript received September 2, 2017; revised November 26, 2017; accepted December 12, 2017. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIP) (No.2016R1A2B4011712). Paper no. TII-17-2066.
(Corresponding author: Sung Wook Baik.)

K. Muhammad, J. Ahmad and S. W. Baik are with the Intelligent Media Laboratory, Digital Contents Research Institute, Sejong University, Seoul 143-747, Republic of Korea (e-mail: khan.muhammad@ieee.org; jamilahmad@ieee.org; sbaik@sejong.ac.kr).

R. Hamza is with the Department of Computer Science, University of Batna 2, Batna 05078, Algeria (e-mail: r.hamza@univ-batna2.dz).

J. Lloret is with the Instituto de Investigacion para la Gestion Integrada de Zonas Costeras, Universitat Politecnica de Valencia, Valenciana 46022, Spain (e-mail: jlloret@dcom.upv.es).

H. Wang is with GoPerception Laboratory, NY, USA and Cornell University, USA (e-mail: hw496@goperception.com).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TII.2018.2791944

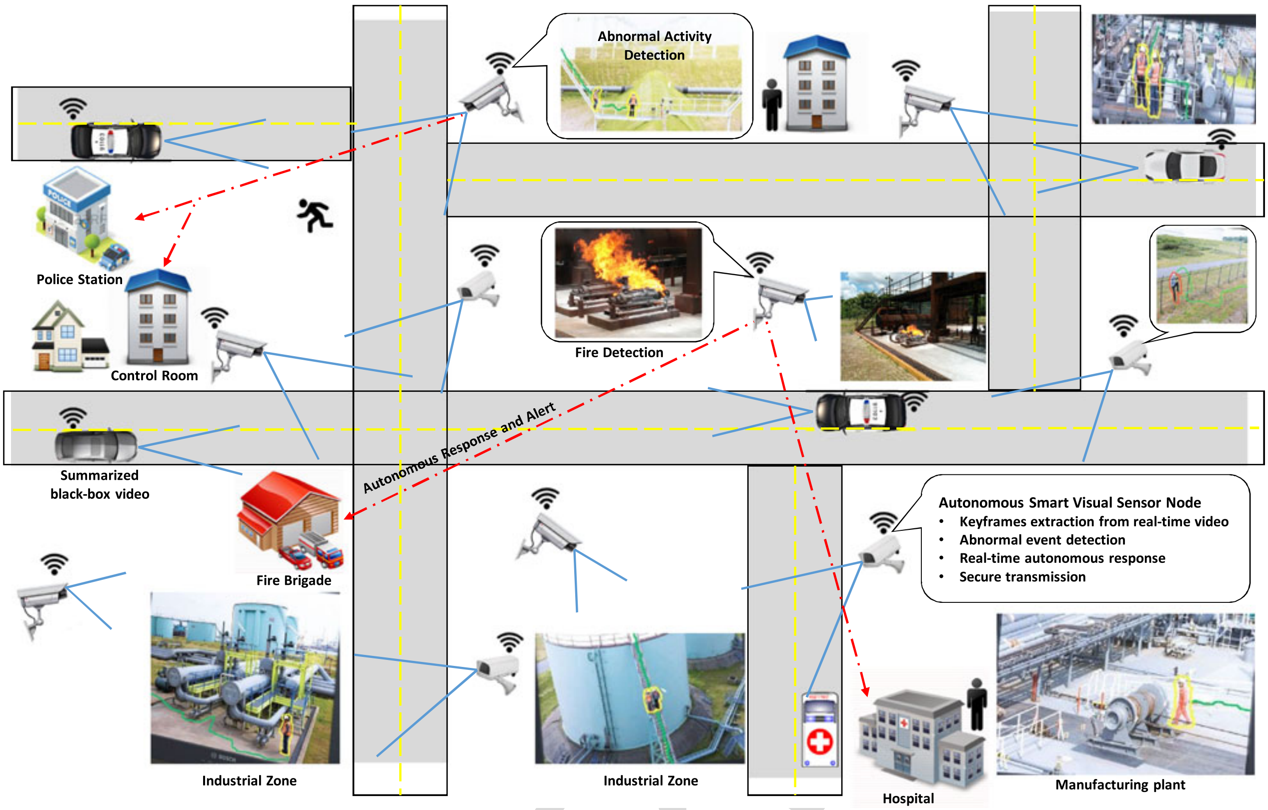

Fig. 1. Smart and secure surveillance framework using IoT infrastructure in industrial environment.

mechanism because any modification to the transmitted data can greatly affect the analyst's decision at the BS. Furthermore, utilization of a dedicated spectrum for transmission of multimedia data in WMSNs is comparatively difficult due to the congested bandwidth allocation mechanism.

Therefore, in this paper, we address these problems by using an intelligent and power-efficient system that can make each sensor node intelligent and autonomous enough to collect only the important data in real time and take the appropriate action accordingly, thus, reducing the bandwidth consumption and transmission cost. Furthermore, we develop a security prototype for secure transmission of semantically relevant visual data to a fusion center with improved spectrum utilization and preservation of the limited resources of WMSNs. Technically, our system uses image encryption to encrypt the visual contents prior to transmission, thus, increasing the security during communication within industrial WMSNs. For encryption of digital images, the commonly used approaches include nonlinear chaotic systems, as verified from the recent literature. For instance, in our previous work [7], we used a Zaslavsky chaotic map without employing finite computations of the pseudo random number generator (PRNG) for symmetric image encryption using permutation and diffusion. Later on, in another work [8], we applied our algorithm to the extracted keyframes of a wireless capsule endoscopy (WCE) procedure using video summarization [9], [10] and proved its ability to withstand all known attacks. This ensured the dissemination of important keyframes to healthcare centers and gastroenterologists for personalized WCE.

In this paper, we propose an energy-friendly image encryption algorithm using one chaotic map employed in PRNG and a cryptosystem structure. Probabilistic cipher is achieved using embedded random bits with plain images, providing randomized ciphered images that are indistinguishable from random noise. Various tests and results show the excellent performance of the proposed cryptosystem, which exceeds several state-of-the-art algorithms. The simulation and security analysis indicate that the proposed encryption algorithm can produce different ciphered images with a high level of security and limited processing time, making it more suitable for industrial IoT systems.

The rest of this paper is organized as follows: Section II demonstrates the proposed system in detail. Section III presents the experimental results, followed by concluding remarks and future directions in Section IV.

II. Proposed Secure Surveillance Framework

The rise in demand for constant surveillance, improvement in visual sensor technologies, and the progress in IoT technologies has necessitated the efficient management and timely analysis of the multimedia big data generated by the ever growing number of surveillance networks in industrial systems. These technologies make it possible to automatically analyze the video data so as to generate a real-time autonomous response. Visual sensor networks have become smarter, with improved storage and processing capabilities enabling them to perform complex data processing in real time. In the case of multiview surveillance videos

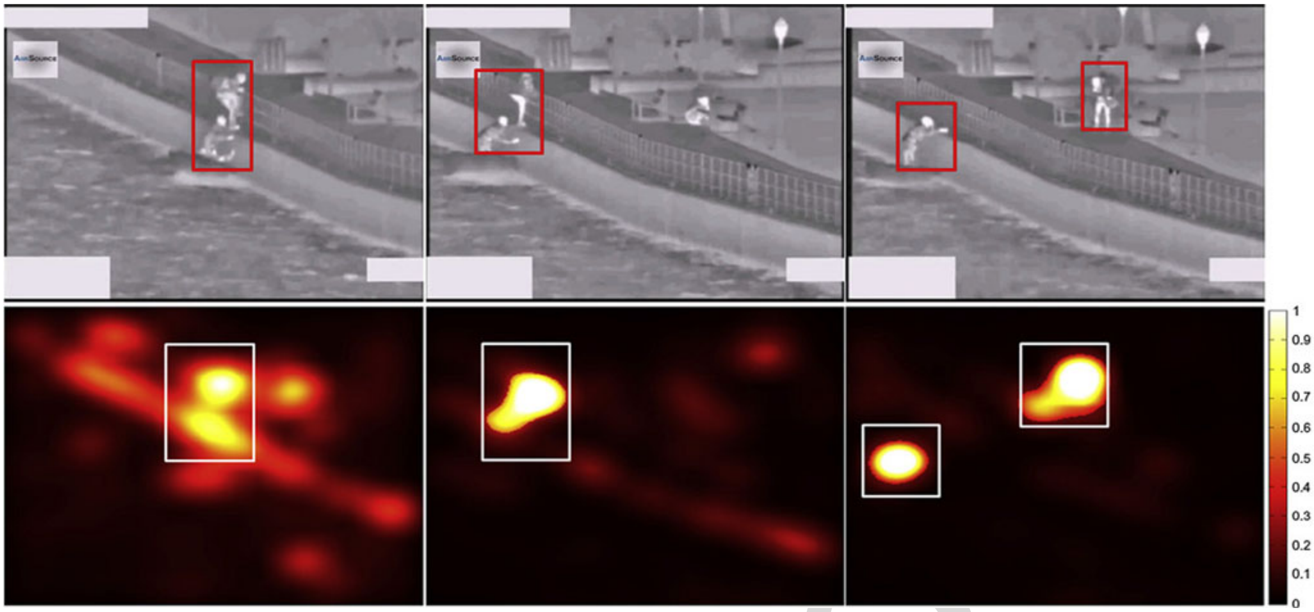

Fig. 2. Illustrating salient motion detection. The first row shows two persons crossing the fence and the second row shows the salient motion of objects detected by our approach.

captured in industrial environments, their processing abilities can be used to analyze the video stream to identify keyframes and then discard irrelevant and redundant visual data, thus minimizing the bandwidth requirements. The improved communication abilities of sensor nodes can be used to collaboratively perform sophisticated scene analysis in order to generate multiview summaries of surveillance videos in real time. The smart sensors can be used to generate an autonomous response after detecting abnormal events, such as fire in industrial zones, by utilizing the IoT infrastructure. Furthermore, the security of the keyframes can be ensured by applying lightweight encryption algorithms, considering the processing capabilities, memory, and transmission constraints. An overview of the proposed system is given in Fig. 1. The details of this framework and its main embodiments are illustrated in the subsequent sections.

A. Keyframes Extraction Using Video Summarization From the Stream of Visual Sensors

The VPH in industrial surveillance networks collects visual data from visual sensors in the form of video frames, resulting in large volumes of video data. Due to the energy and bandwidth constraints of WMSNs, the transmission of all of the streaming data is impractical because of the larger distance between the BS and VPH. To tackle this issue, researchers have employed different compression [11] and video summarization methods [12] to reduce the volume of visual data at the VPH so that only informative video frames are forwarded to the BS for processing. Considering the bandwidth and energy constraints, we employ an energy-friendly keyframe extraction approach from our recent work [4] to reduce the redundancy. Our keyframes extraction algorithm is lightweight because it uses novel integral-image features for salient motion detection. This computationally efficient algorithm can be employed for small devices, such as visual sensors that have energy, processing, and bandwidth constraints. This is evident from [13], where the authors experimentally proved that the results of integral images are 15 times faster than existing methods of object detection. To extract keyframes using this approach, first, the integral image is computed for each frame captured by the visual camera, then, background bootstrapping is conducted, which is essential for the removal of background motion and accurate estimation of salient motion. Salient motion can be measured by computing the changes in image block values in neighboring frames. It is robust to even small background motion, as it uses background model and integral image based temporal gradients for salient motion. This can be verified from Fig. 2, where the salient motion detection by our scheme is illustrated using a few frames from a sample video of an illegal border crossing.

In the given sample video, there is significant motion clutter due to the strong wind and river waves that continuously change the background pattern, thus, making the salient motion detection more challenging. Despite these challenges, this approach detects the salient motion correctly, as shown in Fig. 2. Based on the salient motion detection, the informative frames are selected and then passed to the encryption module for lightweight encryption.

B. Probabilistic Keyframes Encryption Algorithm

This section illustrates the encryption process for the keyframes extracted from the stream of visual sensors in an IoT industrial environment. The proposed algorithm has two major components: The first component aims to use a recent two-dimensional (2-D) chaotic map [14] to produce PRNG suitable for our proposed image encryption, and the second component executes one round of permutation–diffusion processes for the keyframe under consideration. Most surveillance systems

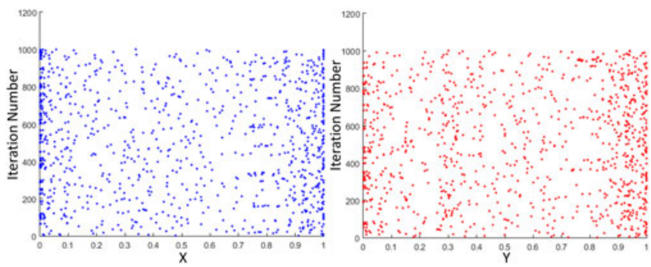

Fig. 3. Plot distributions of (*x*, *y*) chaotic sequence.

capture videos in RGB format through visual sensors with a high resolution. Thus, we propose a fast RGB image encryption algorithm that guarantees the privacy as well as the confidentiality of the keyframes. Furthermore, we use a randomized approach, making it infeasible for attackers to learn anything about the original data from the ciphered frames. This restricts the availability to attackers of the information required to build a cryptanalysis model.

1) 2-D Logistic-Sine System: A 2-D logistic-adjusted-sine map (LASM) is presented with efficiencies and high sensitivity to initial values and a complex chaotic behavior of its generated sequences. The mathematical equation of the LASM is as follows:

$$\begin{cases} x_{i+1} = \sin\left(\pi u \left(y_i + 3\right) x_i \left(1 - x_i\right)\right) \\ y_{i+1} = \sin\left(\pi u \left(x_i + 3\right) y_i \left(1 - y_i\right)\right) \end{cases}. \quad (1)$$

Herein, all values $(x,\ y,\ u)$ are within [0, 1]. The properties of this map have important features, such as ergodicity, unpredictability, and sensitivity to initial values [14]. Fig. 3(a) and (b) shows the plot of sequence values generated directly from the LASM. As shown in Fig. 3, this map has good uniform distribution for its sequences with complex chaotic behaviors and better unpredictability [14]. We chose this map to design our PRNG and employed it in our image encryption scheme.

We design a new PRNG based on the LASM, whose secret keys are used to generate the chaotic numbers sequence related to the size of the plain image. In addition, we use the aggregate of plain image pixels to guarantee a high level of security against all chosen attacks. The procedure of generating chaotic sequences using the LASM is shown in Algorithm 1.

Herein, we compute the sum of the pixels of the keyframe or the input sequence so that the generated sequences are related to the original keyframe. To get rid of the effect of the initial values, we remove the first three numbers generated from the sequence. For ease of understanding, we denote the pseudorandom number generator in Algorithm 1 by PRNG, where the inputs are a set of numbers of secret keys and a sequence of numbers.

2) Keyframe Encryption: The major steps of encrypting a keyframe are described in this section. First, we set the initial values $x_0, y_0, u_0, x_1, y_1, u_1$ as secret keys to make exhaustive attacks ineffective and useless. Coding the pixels of the keyframe starts with embedding true chaotic bits into only one channel of the original keyframe. Then, confusion and diffusion operations are designed to randomly change the pixel values and shuffle the pixel positions, respectively. Since real-time applications need a

Algorithm 1: Generation of Chaotic Sequences Using LASM (PRNG).

Input: $(x_0,\ y_0,\ u,\ P)$
1: $[a, b, c] \leftarrow \text{size}(P)$
2: Sum $= \sum_i \sum_j P.$
3: IF Sum $=\ 0$
  $S \leftarrow 0;$
 Else
  $S0 = 2 + \text{abs}\ (\log 10\ (\text{sum}^{-1}))$
  $S = e^{(S0)} \times \text{Sum}^{-1g}$
End
4: $x = x_0 + S;\ y = y_0 + S;\ u = u + S$
5: Sequence $\leftarrow$ zeros$(a \times b \times c, 1)$
6: For $i = 1$ *to* ceil$((a \times b \times c)/2)$
  $x_{i+1} = \sin(\pi\, u(y_i + 3)x_i(1 - x_i))$
  $y_{i+1} = \sin(\pi\, u(x_i + 3)y_i(1 - y_i))$
  Sequence$(2i)$ = floor$(10^{10} \times x_{i+1})$ mod 256
  Sequence$(2i + 1)$ = floor$(10^{10} \times y_{i+1})$ mod 256
 End
 Output: Sequence

fast algorithm, we thus minimize the steps and computations in our encryption scheme to comply with the real-time processing demands of IoT devices in industrial zones. It should be noted that our proposed method can encrypt images of all dimensions with size [*a*, *b*, 3], where "*a*" and "*b*" are integer numbers.

Fig. 4 shows the visual encryption and decryption for a selected keyframe from the surveillance streams. The steps of the encryption are highlighted as follows.

Step 1: Let the keyframe be denoted by I of size $[a \times b \times 3]$. First, the chaotic sequences of numbers are constructed as described in Algorithm 1. The generated sequence is denoted by P_1 as follows:

$$P_1 = \text{prsg}\left(x_0,\ y_0,\ u_0, 0\right) \quad (2)$$

Herein, we set zeros with same size as the plain keyframe I instead of the plain image, so that $S = 0$, as given in Algorithm 1.

Step 2: Next, we apply the initial processing as follows:

$$[I_R\ I_G I_B] \leftarrow I$$
$$C_R = \text{LSBNoise}\left(I_R\right) \oplus I_G \oplus I_B$$
$$C_G = C_R \oplus I_B$$
$$C_B = C_R \oplus I_G$$

$C_1 \leftarrow [C_R\ C_G\ C_B]$, reshape the three matrices $(C_R\ C_G\ C_B)$ into the 1-D vector C_1

$$C_{\text{initial}} = C_1 \oplus P_1.$$

Here, LSBNoise uses a random noise bit at the position of the least significant bit (LSB). It consists of the integration of the probabilistic sound encryption LSB [15]. In this step, we use a random source to ensure that each produced bit has the possibility of 50%. Next, we generate a random bits matrix with

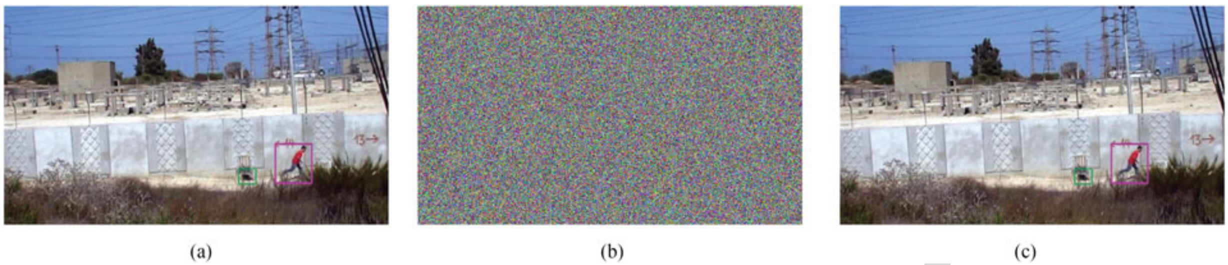

Fig. 4. Illustrating encryption/decryption using a sample frame from surveillance of interest.

size $[a, b]$, followed by embedding the random bits in the plain image using an XOR operation.

It should be noted that the proposed image encryption can encrypt both grayscale and color images without any issue. For a grayscale image, we treat its matrix as a red channel only and embed the noise bits in the entire grayscale matrix, followed by the rest of the encryption steps. For color images, we reshape the image matrices into a 1-D vector, i.e., $[1, 3{*}w{*}h]$. The inverse operation is possible, which restores the same number of matrices at the final stage of encryption. Thus, a grayscale image with one matrix or an RGB image with three matrices will not disturb our cryptosystem.

Step 3: We generate two sequences P_2 and P_3, respectively, as follows:

$$\begin{cases} P_2 = P_1 \oplus \text{prsg}\,(x_1,\ y_1,\ u_1, C_{\text{initial}}) \\ P_3 = \text{prsg}\,(x_0,\ y_0,\ u_0, C_{\text{initial}}) \end{cases}. \tag{3}$$

Note: The total number of pixels in the original keyframe is defined as $a \times b \times c$. Therefore, all the generated sequences from Algorithm 1 must be of the same size.

Step 4: Next, we sort the sequences P_2 and P_3 in ascending order to obtain the indices sequences π and π' as shown in (4) and (5). Thus, the generated sequences represent permutation matrices

$$\text{Sort}\,(P_1)\ = P_1' = \begin{bmatrix} 1 & 2 & 3 & & a \times b \times c \\ & , & , & \ldots, & \\ \pi_1 & \pi_2 & \pi_3 & & \pi_{a \times b \times c} \end{bmatrix} \tag{4}$$

$$\text{Sort}\,(P_2)\ = P_2' = \begin{bmatrix} 1 & 2 & 3 & & a \times b \times c \\ & , & , & \ldots, & \\ \pi'_1 & \pi'_2 & \pi'_3 & & \pi'_{a \times b \times c} \end{bmatrix}. \tag{5}$$

Step 5: Next, we shuffle C using the sort index of the new sequences. Here, we employ the P-box of P_2' followed by the P-box of P_3'.

Step 6: Next, we shuffle C using the P-box of P_3', followed by the P-box of P_2'.

Step 7: Finally, we reshape the obtained matrix of the previous steps into three matrices corresponding to the RGB matrices. The obtained matrix is denoted by "C," which is the ciphered frame for plain image I.

3) Keyframe Decryption: The decoding process is the inverse of the encryption mechanism, aiming to get the original keyframe. The following steps are used to restore the original keyframe from the encrypted frame using the exact values of the secret keys.

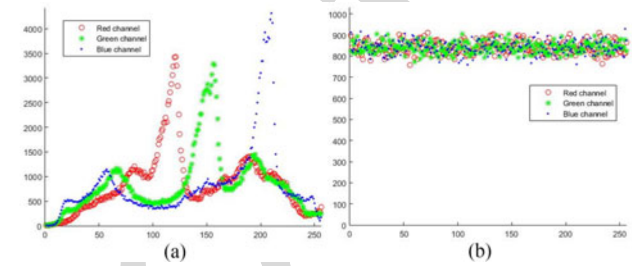

Fig. 5. (a) Histogram of the individual plane of an RGB keyframe given in Fig. 4(a); (b) histogram of the three planes for the encrypted keyframe given in Fig. 4(b).

Step 1: Read the ciphered keyframe C_{initial} and get its size $[a, b]$.

Step 2: Reshape the image matrices into one matrix with size $[a, 3, b]$.

Step 3: Generate the chaotic sequences $P_1,\ P_2,$ and P_3 using Algorithm 1 as follows:

$$\begin{cases} P_1 = \text{prsg}\,(x_0,\ y_0,\ u_0, 0) \\ P_2 = P_1 \oplus \text{prsg}\,(\,x_1,\ y_1,\ u_1, C_{\text{initial}})\,. \\ P_3 = \text{prsg}\,(x_0,\ y_0,\ u_0, C_{\text{initial}}) \end{cases} \tag{6}$$

Step 4: Use the bijection property of the permutation matrix of P_2' and P_3' to restore the original position of the pixels. For this, first we use the inverse P-box of P_3' followed by the inverse P-box of P_2'.

Step 5: Repeat step 4 by changing the order of the P-box, i.e., use the inverse P-box of P_2' first, followed by using the inverse P-box of P_3'. The obtained matrix is denoted by D_4.

Step 6: Apply the final processing steps as follows:

$D_{\text{Final}} = D_4 \oplus P_1$, Reshape the obtained matrix into three matrices $D_R'\ D_G'\ D_B'$ corresponding to the RGB matrix.

$$D_R \leftarrow D_R' \oplus D_G' \oplus D_B'\ ,\ D_G \leftarrow D_G' \oplus D_R',\ \text{and}$$
$$D_B \leftarrow D_B' \oplus D_R'.$$

Step 7: The obtained matrix, denoted by "D," consists of D_R, D_G, and D_B matrices, indicating the decrypted keyframe.

III. Experimental Results and Discussion

This section illustrates the performance evaluation of the proposed system from different perspectives. We used MATLAB R2015a in the Windows 10 environment with an i7 processor of 2.4 GHz and 12 GB of RAM for the experimentation,

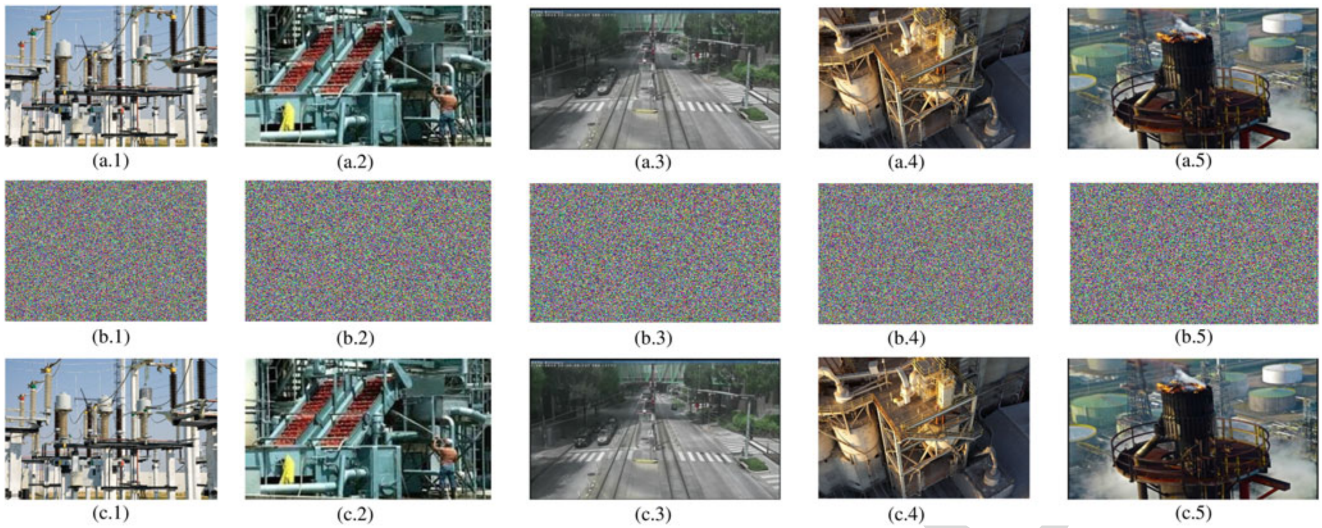

Fig. 6. (a.i) Keyframes, (b.i) encrypted keyframes, and (c.i) decrypted keyframes, respectively (from left to right, and ($i \in \{1, 2, 3, 4, 5\}$).

TABLE I
INFORMATION ENTROPY TESTS

Name	Size	Keyframe			Ciphered		
		R	G	B	R	G	B
Zeros pixel	[1024 1024][3]	0	0	0	7.9998	7.9998	7.9998
Keyframe 1	[240 352] [3]	6.6640	6.6580	6.7605	7.9976	7.9976	7.9976
Keyframe 2	[240 352] [3]	6.2363	6.0248	5.9998	7.9981	7.9978	7.9979
Keyframe 3	[240 352] [3]	7.7660	7.6599	7.7855	7.9975	7.9977	7.9979
Keyframe 4	[240 352] [3]	6.8212	6.7584	6.7003	7.9979	7.9975	7.9979
Keyframe 5	[240 352] [3]	6.8679	6.8531	6.7077	7.9979	7.9976	7.9978
Keyframe 6	[240 352] [3]	6.4410	6.3789	6.4770	7.9978	7.9978	7.9979

simulation, and analysis. We set 0.67 0.9 0.4 0.67 0.9 0.4 as a default secret key for the proposed image encryption during the experimental tests.

A. Visual and Histogram Tests

The histogram of an image describes its pixels distribution by plotting the number of pixels at each color intensity level [16]. Fig. 5 shows the histogram of a plain image and encrypted image before and after the encryption in three components R, G, and B, respectively. The histograms in the three components of the encrypted image are very uniform and completely different from the histograms of the plain image.

Fig. 6 shows different keyframes and their encrypted and decrypted versions extracted from visual data of surveillance in industrial networks. Thus, our proposed image encryption algorithm can withstand the statistical attacks.

B. Information Entropy

It is agreed in the image encryption community that the ciphered images should appear as truly random sources. To verify this, information entropy is the most important metric that decides whether the sources are random or not. We calculate the entropy of an image (the entropy of a source) with $P(c_i)$ representing the probability of a pixel, using the following equation:

$$S(C) = -\sum_{i=1}^{255} P(c_i)\ \log_2 P(c_i). \tag{7}$$

According to this test, the information entropy of the ciphered keyframe should be close to 8. Table I shows the numerical values of the entropy for a set of keyframes and their corresponding ciphered keyframes for three individual channels. All the values obtained from Table I are close to 8. Therefore, our proposed image encryption produces a secure ciphered image with a randomlike source.

C. NPCR and UACI

In this section, we employ the number of changing pixel rate (NCPR) and the unified averaged changed intensity (UACI) tests [17] to prove that our proposed cryptosystem can avoid differential attacks against ciphered data. Basically, the attacker aims to cipher two images, differing in a pixel, and look at the difference between the corresponding ciphered images. Here, the difference between the ciphered data should not show any black-zone blocks. In this regard, we produce two ciphered images generated from our proposed image encryption. We investigated the ability to resist the differential attacks with the propriety of probabilistic encryption. Here, we tested the NPCR and UACI

TABLE II
NPCR AND UACI TESTS RESULTS FOR EACH CHANNEL OF RGB

	Keyframe1		Keyframe2		Keyframe3		Keyframe4		Keyframe5	
	NPCR	UACI	NPCR	UACI	NPCR	UACI	NPCR	UACI	NPCR	UACI
R	99.5881	33.3848	99.5713	33.3379	99.6070	33.4251	99.6009	33.4910	99.6165	33.3546
G	99.6283	33.4955	99.6123	33.5213	99.5608	33.4013	99.5899	33.3394	99.6094	33.3943
B	99.5999	33.4559	99.6059	33.4705	99.6307	33.5713	99.6311	33.4804	99.6046	33.4404

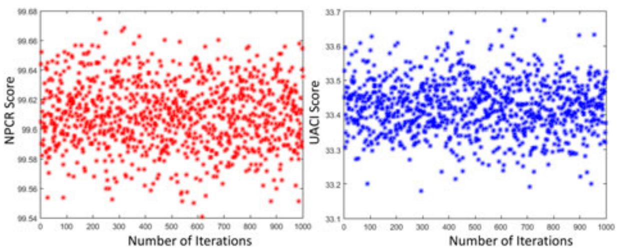

Fig. 7. Evaluation of the probabilistic image encryption using NPCR and UACI tests for 1000 repeats.

TABLE III
COMPARISON RESULTS FOR EACH CHANNEL OF RGB

	Our	Belazi *et al.* [19]	Wei *et al.* [20]	Zhou *et al.* [21]	Zhou *et al.* [22]
NPCR	99.6125	99.6177	99.2172	99.60	99.6098
UACI	33.4451	33.6694	33.4058	33.40	33.4384

scores of two ciphered images C1 and C2 that are generated from the same image I using the same secret keys. Equations (8) and (9) present the formulas of these tests as follows:

$$\mathrm{NPCR}\left(C_1, C_2\right) = \sum_{i,j} \frac{S\left(i,j\right)}{D} \times 100\ \% \tag{8}$$

$$\mathrm{UACI}\left(C_1, C_2\right) = \sum_{i,j} \frac{C_1\left(i,j\right) - C_2\left(i,j\right)}{255 \times D} \times 100\ \%. \tag{9}$$

Herein, "D" denotes the number of pixels and "S" is represented by

$$S\left(i,j\right) = \begin{cases} 0, & \text{if } C_1\left(i,j\right) = C_2\left(i,j\right) \\ 1, & \text{Elsewise.} \end{cases} \tag{10}$$

Our proposed image encryption is a randomized algorithm, which produces completely different encrypted images for the same plain image using the same secret key. We submitted both ciphered images C1 and C2 to the NPCR and UACI tests and collected the results for a set of images as listed in Table II. The results demonstrate that our cryptosystem is semantically secure and can ensure that the attacker cannot find any information between the ciphered images and the original ones. The results prove that each encryption is completely different from the next (randomly ciphered). Fig. 7 shows the results of the NPCR and UACI test repeated 1000 times for zero pixels with size [256, 256], [3], where we took the average result for the three plans (RGB).

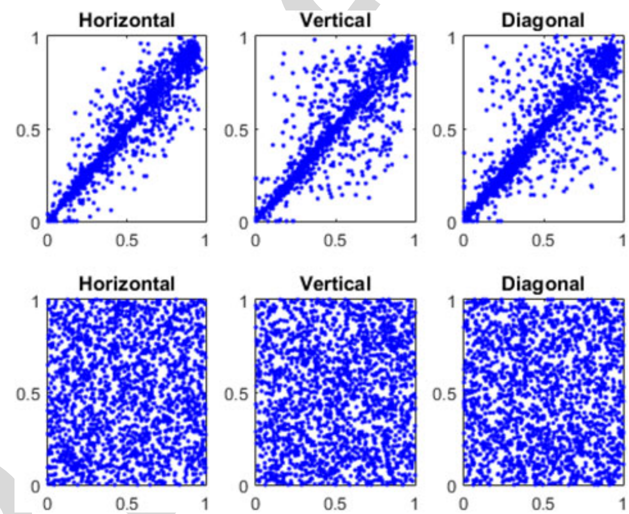

Fig. 8. Distribution of two adjacent pixels in the plain and encrypted image in the blue channel over horizontal, vertical, and diagonal directions.

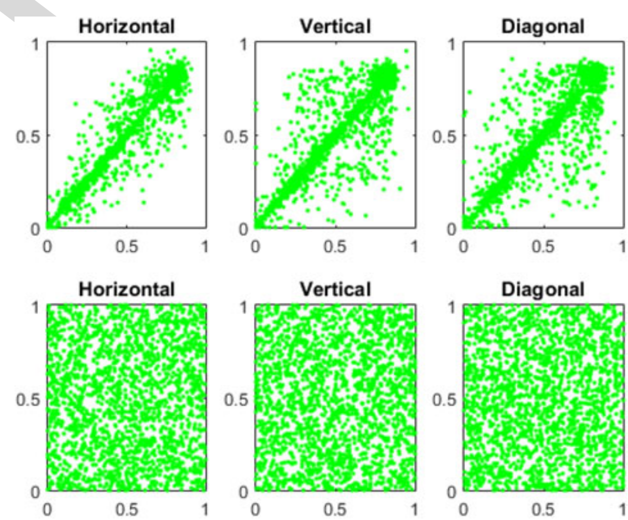

Fig. 9. Distribution of two adjacent pixels in the plain and encrypted image in the green channel over horizontal, vertical, and diagonal directions.

Our proposed scheme successfully passed these tests and exceeded all theoretical values [7]. In addition, we compared the performance of our algorithm with other recent encryption algorithms in Table III, and can demonstrate the effectiveness of our proposed scheme compared with other methods. All the results demonstrated that our proposed image encryption has a strong ability to resist differential attacks.

TABLE IV
CC OF ADJACENT PIXELS TESTS

	Component	Keyframe			Ciphered		
		Horizontal	Vertical	Diagonal	Horizontal	Vertical	Diagonal
Keyframe1	R	0.9716	0.8707	0.8569	0.0035	0.0055	8.034e−04
	G	0.9660	0.8459	0.8288	− 0.0026	− 0.0044	0.0016
	B	0.9663	0.8464	0.8292	0.0025	− 3.594e−04	0.0034
Keyframe2	R	0.9860	0.9442	0.9304	− 0.0014	− 0.0042	0.0092
	G	0.9862	0.9434	0.9296	− 0.0034	− 0.0033	− 0.0024
	B	0.9872	0.9491	0.9364	0.0077	− 0.0029	0.0017
Keyframe3	R	0.9376	0.8672	0.8470	0.0030	0.0075	− 0.0053
	G	0.9382	0.8691	0.8494	0.0063	− 0.0024	− 0.0051
	B	0.9469	0.8881	0.8711	0.0017	− 0.0023	− 0.0030
Keyframe4	R	0.9948	0.9908	0.9884	− 0.0010	− 0.0022	0.0012
	G	0.9919	0.9852	0.9819	− 0.0015	0.0016	− 0.0017
	B	0.9911	0.9836	0.9800	0.0025	− 0.0004	0.0003

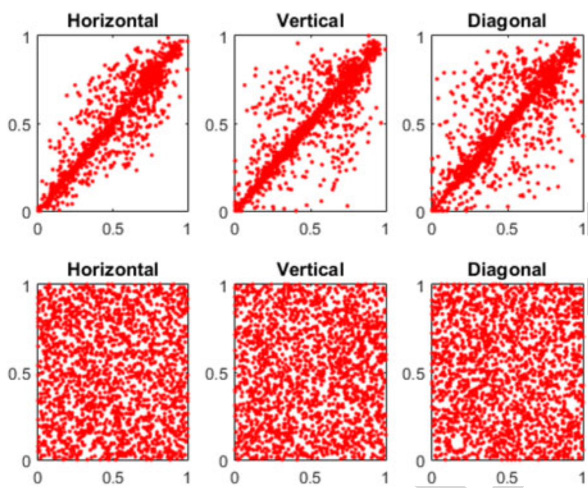

Fig. 10. Distribution of two adjacent pixels in the plain and encrypted image in the red channel over horizontal, vertical, and diagonal directions.

D. Correlations Analysis

A plain image has high information redundancy and high correlations with its neighboring pixels. Generally speaking, an original image has a correlation coefficient (CC) almost equal to 1. Therefore, image encryption should be able to eliminate these correlations, indicating that the ideal value of an encrypted image is CC $= 0$ [18]. The correlation of two adjacent pixels is presented mathematically as follows:

$$\mathrm{CC}_{xy} = \frac{\mathrm{cov}\left(x, y\right)}{\sqrt{D\left(x\right) \times D\left(y\right)}} \tag{11}$$

$$\mathrm{cov}(x, y) = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - E(x)\right)\left(y_i - E(y)\right) \tag{12}$$

$$D\left(x\right) = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - E(x)\right)^2 \tag{13}$$

$$E\left(x\right) = \frac{1}{n}\sum_{i=1}^{n} x_i. \tag{14}$$

TABLE V
COMPARISON OF CC OF ADJACENT PIXELS TESTS

Algorithm	Our	[24]	[19]	[25]	[23]
CC score	0.0034	0.0060	0.0129	0.0031	0.0722

TABLE VI
KEY SPACE COMPARISON

Algorithm	Our	[24]	[23]	[25]
Space key	10^{90}	0.25×10^{64}	10^{56}	2^{180}

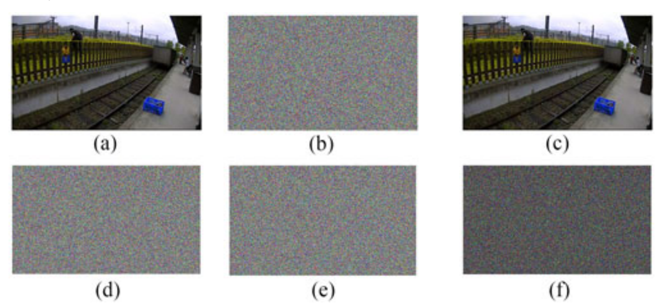

Fig. 11. (a) Plain keyframe, (b) encrypted keyframe using the secret key 0.67 0.9 0.4 0.67 0.9 0.4; (c) decrypted keyframe using the secret key 0.67 0.9 0.4 0.67 0.9 0.4; (d) decrypted keyframe using the secret key $0.67 + 10^{-15}$ 0.9 0.4 0.67 0.9 0.4; (e) decrypted keyframe using the secret key 0.67 $0.9 + 10^{-15}$ 0.4 0.67 0.9 0.4; (f) difference image between (d) and (e).

TABLE VII
ENCRYPTION/DECRYPTION SPEED TEST RESULTS

Size of keyframe	[256, 256, 3]	[512, 512, 3]	[1024, 1024, 3]	[2048, 2048, 3]
Speed (s)	0.1616	0.6708	2.821	11.5471

TABLE VIII
COMPARISON RESULTS BETWEEN OUR ALGORITHM AND PREVIOUS RECENT SCHEMES

	Image size	Key space	Speed (ms)	Entropy	CCa	NPCR	UACI
Our	[1024, 1024], [3]	10^{90}	2821	7.9998	0.0035	99.6125	33.4451
[19]	[1024,1024], [1]	2^{624}	2513	7.9998	0.0129	99.6177	33.6694
[24]	[256,256], [1]	0.25×10^{64}	1320	7.9974	0.0060	99.6200	33.5100
[23]	[256,256], [1]	10^{56}	547	7.9959	0.0722	>99	≅33.43

We employed the statistical test of correlation of two adjacent pixels in encrypted keyframes. We randomly select 2000 pixels in keyframes and their corresponding adjacent pixels in each channel from the RGB space along with the horizontal, vertical, and diagonal directions. Figs. 8–10 show the visual results for the distributions of two adjacent pixels in a keyframe and the corresponding ciphered keyframe in the blue, green, and red channels over the horizontal, vertical, and diagonal directions. The graphs in the first row are for the plain keyframe, whereas the graphs in the second row are for the encrypted keyframe. It can be noted that the plots vary greatly in both the original keyframe and the encrypted keyframe. The dots are well distributed with a good uniform probability distribution in the plot of the ciphered keyframe. Dots are located on the diagonal line in the plot of the original keyframe.

Next, we used the selected pixels of keyframes and their corresponding encrypted keyframes to compute the numerical scores of CC in the three channels along the horizontal, vertical, and diagonal directions. Table IV shows the results of this test with different sets of keyframes and their ciphered versions with numerical values near to one and zero, respectively. Finally, we compared the average of the numerical results with the scores of other recent methods [19], [23], [24]. The results show that our proposed algorithm achieves comparable or better scores, as reported in Table V. Thus, our proposed image encryption can considerably reduce the inherent correlation of the original adjacent pixels.

E. Analysis of Secret Key

To resist exhaustive attacks, the space key of an encryption algorithm should be at least 2^{128}. In our proposed image encryption, we set $(x_0,\ y_0,\ u_0,\ x_1,\ y_1,\ u_1)$ as secret keys. The space key in our work can be computed with more than 10^{90} and, with such a large space key, there is no need for brute force to break our proposed image encryption. Moreover, the space key is larger than other recent schemes, as shown in Table VI.

Since our proposed image encryption is probabilistic, the ciphered image will accordingly change completely for each encryption using the same keyframe and secret keys. Therefore, our proposed image encryption does not give any useful information to attackers, thus validating its security. Fig. 11 shows that decryption is an option only with the exact secret keys, and that our proposed cryptosystem is robust against differential attacks at decryption processes. Therefore, our algorithm is highly sensitive to the secret key and provides a high level of security for the keyframes.

F. Speed Tests and Performances Comparison

In this part, we show the results of the encryption/ decryption execution time test for a set of keyframes with different sizes. Table VII shows the numerical value obtained after encrypting the keyframes. In our proposed encryption scheme, the encryption and decryption have the same execution time. As shown in Table VII, the running time of the proposed scheme is fast, making it more suitable for real-time applications, such as secure surveillance.

In addition, we compared the performance of our proposed image encryption with other recent encryption schemes [19], [23], [24]. Table VIII shows the comparison between our proposed method and these other cryptosystems. It is clear that the results obtained from our algorithm exceeded the ideal values for these tests [7] and are comparable to other algorithms. All these schemes have reported a good score and present a secure level of confidentiality for the images. Our CC average (CCa) score is obtained from the average of all values of the CC. As shown, CCa in our algorithm has the lowest values, which reflect the strength of the proposed algorithm for eliminating the strong correlation of adjacent pixels of the plain image. Since we compared our performance with a different set of images under various platforms and system characteristics with many factors, we can only approximate the faster algorithm. Our proposed image encryption has a good execution rate of 1310.7 kb/s. The work in [24] has 49.64 kb/s, [19] has 0.4173 kb/s, and [23] has 0.1198 kb/s. These statistics indicate that the execution time of our algorithm is better than the other mentioned algorithms.

IV. CONCLUSION AND FUTURE WORK

Due to recent advances in IoT-assisted networks for surveillance in industrial environments, a significant amount of redundant video data are generated. Its transmission, analysis, and management are difficult and challenging, requiring image prioritization. In this paper, an efficient video summarization method is first used to extract the informative frames from the surveillance video data, which can be used for abnormal event detection. Since the extracted keyframes are important for further analysis, their privacy and security is of paramount importance during transmission. Therefore, we proposed a fast probabilistic and lightweight algorithm for the encryption of keyframes prior to transmission, considering the memory and processing requirements of constrained devices, which increase its suitability for industrial IoT systems. Our algorithm is secure because an attacker cannot collect any useful information about a keyframe from its corresponding ciphered image. The experimental results verify the efficiency, security, and robustness of our algorithm compared to other image encryption methods. Furthermore, it also confirms its effectiveness for reducing the bandwidth, storage, and transmission cost, as well as reducing the browsing time of analysts dealing with large volumes of surveillance data to make decisions about abnormal events, such as suspicious activity detection and fire detection in industrial environments.

This paper mainly focuses on video data of visual sensors and does not consider data collected in the IoT environment from other types of sensors. Further research can be conducted to incorporate data from other diverse devices for numerous applications [26]–[29] and further improve the security measures in other specific areas [30]–[32]. Another research direction is to use dynamic keys instead of traditional encryption keys to further improve the security of the overall framework.

References

[1] L. Da Xu, W. He, and S. Li, "Internet of things in industries: A survey," *IEEE Trans. Ind. Informat.*, vol. 10, no. 4, pp. 2233–2243, Nov. 2014.

[2] M. Sajjad, I. Mehmood, and S. W. Baik, "Sparse representations-based super-resolution of key-frames extracted from frames-sequences generated by a visual sensor network," *Sensors*, vol. 14, pp. 3652–3674, 2014.

[3] J. Lloret, I. Bosch, S. Sendra, and A. Serrano, "A wireless sensor network for vineyard monitoring that uses image processing," *Sensors*, vol. 11, pp. 6165–6196, 2011.

[4] I. Mehmood, M. Sajjad, W. Ejaz, and S. W. Baik, "Saliency-directed prioritization of visual data in wireless surveillance networks," *Inf. Fusion*, vol. 24, pp. 16–30, 2015.

[5] Q. Wu, M. Tao, D. W. K. Ng, W. Chen, and R. Schober, "Energy-efficient resource allocation for wireless powered communication networks," *IEEE Trans. Wireless Commun.*, vol. 15, no. 3, pp. 2312–2327, Mar. 2016.

[6] D. Zhang, G. Li, K. Zheng, X. Ming, and Z.-H. Pan, "An energy-balanced routing method based on forward-aware factor for wireless sensor networks," *IEEE Trans. Ind. Informat.*, vol. 10, no. 1, pp. 766–773, Feb. 2014.

[7] R. Hamza and F. Titouna, "A novel sensitive image encryption algorithm based on the Zaslavsky chaotic map," *Inf. Secur. J., Global Perspective*, vol. 25, pp. 162–179, Dec. 1, 2016.

[8] R. Hamza, K. Muhammad, Z. Lv, and F. Titouna, "Secure video summarization framework for personalized wireless capsule endoscopy," *Pervasive Mobile Comput.*, vol. 41, pp. 436–450, 2017.

[9] K. Muhammad, M. Sajjad, M. Y. Lee, and S. W. Baik, "Efficient visual attention driven framework for key frames extraction from hysteroscopy videos," *Biomed. Signal Process. Control*, vol. 33, pp. 161–168, 2017.

[10] K. Muhammad, M. Sajjad, and S. W. Baik, "Dual-level security based cyclic18 steganographic method and its application for secure transmission of keyframes during wireless capsule endoscopy," *J. Med. Syst.*, vol. 40, pp. 1–16, 2016.

[11] J. Wu, B. Cheng, M. Wang, and J. Chen, "Energy-aware concurrent multipath transfer for real-time video streaming over heterogeneous wireless networks," *IEEE Trans. Circuits Syst. Video Technol.*,doi: 10.1109/TCSVT.2017.2695368.

[12] R. Panda and A. R. Chowdhury, "Multi-view surveillance video summarization via joint embedding and sparse optimization," *IEEE Trans. Multimedia*, vol. 19, no. 9, pp. 2010–2021, 2017.

[13] P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," in *Proc. 2001 IEEE Comput. Soc. Conf. Comput. Vision Pattern Recognit.*, 2001, pp. I-511–I-518.

[14] Z. Hua and Y. Zhou, "Image encryption using 2d logistic-adjusted-sine map," *Inf. Sci.*, vol. 339, pp. 237–253, 2016.

[15] M. Machkour, A. Saaidi, and M. Benmaati, "A novel image encryption algorithm based on the two-dimensional logistic map and the Latin square image cipher," *3D Res.*, vol. 6, pp. 1–18, 2015.

[16] X. Wu, D. Wang, J. Kurths, and H. Kan, "A novel lossless color image encryption scheme using 2D DWT and 6D hyperchaotic system," *Inf. Sci.*, vol. 349, pp. 137–153, 2016.

[17] Y. Wu, J. P. Noonan, and S. Agaian, "NPCR and UACI randomness tests for image encryption," *Cyber J.: Multidisciplinary J. Sci. Technol., J. Select. Areas Telecommun.*, pp. 31–38, 2011.

[18] B. Norouzi, S. Mirzakuchaki, S. M. Seyedzadeh, and M. R. Mosavi, "A simple, sensitive and secure image encryption algorithm based on hyper-chaotic system with only one round diffusion process," *Multimedia Tools Appl.*, vol. 71, pp. 1469–1497, 2014.

[19] A. Belazi, A. A. A. El-Latif, and S. Belghith, "A novel image encryption scheme based on substitution-permutation network and chaos," *Signal Process.*, vol. 128, pp. 155–170, 2016.

[20] X. Wei, L. Guo, Q. Zhang, J. Zhang, and S. Lian, "A novel color image encryption algorithm based on DNA sequence operation and hyper-chaotic system," *J. Syst. Softw.*, vol. 85, pp. 290–299, 2012.

[21] S. Zhou, Z. Wei, B. Wang, X. Zheng, C. Zhou, and Q. Zhang, "Encryption method based on a new secret key algorithm for color images," *AEU—Int. J. Electron. Commun.*, vol. 70, pp. 1–7, Jan. 2016.

[22] Y. Zhou, Z. Hua, C.-M. Pun, and C. P. Chen, "Cascade chaotic system with applications," *IEEE Trans. Cybern.*, vol. 45, no. 9, pp. 2001–2012, Sep. 2015.

[23] X. Huang, "Image encryption algorithm using chaotic Chebyshev generator," *Nonlinear Dyn.*, vol. 67, pp. 2411–2417, 2012.

[24] L. Xu, Z. Li, J. Li, and W. Hua, "A novel bit-level image encryption algorithm based on chaotic maps," *Opt. Lasers Eng.*, vol. 78, pp. 17–25, 2016.

[25] A. Akhavan, A. Samsudin, and A. Akhshani, "A symmetric image encryption scheme based on combination of nonlinear chaotic maps," *J. Franklin Inst.*, vol. 348, pp. 1797–1813, 2011.

[26] M. Taha, L. Parra, L. Garcia, and J. Lloret, "An Intelligent handover process algorithm in 5G networks: The use case of mobile cameras for environmental surveillance," in *Proc. 2017 IEEE Int. Conf. Commun. Workshops*, 2017, pp. 840–844.

[27] Y. Liu, L. Nie, L. Han, L. Zhang, and D. S. Rosenblum, "Action2Activity: Recognizing complex activities from sensor data," in *Proc. Int. Joint Conf. Artif. Intell.*, 2015, pp. 1617–1623.

[28] Y. Liu, L. Nie, L. Liu, and D. S. Rosenblum, "From action to activity: Sensor-based activity recognition," *Neurocomputing*, vol. 181, pp. 108–115, 2016.

[29] K. Muhammad, M. Sajjad, I. Mehmood, S. Rho, and S. W. Baik, "Image steganography using uncorrelated color space and its application for security of visual contents in online social networks," *Future Gener. Comput. Syst.*, doi: 10.1016/j.future.2016.11.029.

[30] J. Lloret, P. V. Mauri, J. M. Jimenez, and J. R. Diaz, "802.11 g WLANs design for rural environments video-surveillance," in *Proc. Int. Conf. Digit. Telecommun.*, 2006. [Online]. Available: http://ieeexplore.ieee.org/document/1698470/

[31] K. Muhammad, M. Sajjad, I. Mehmood, S. Rho, and S. W. Baik, "A novel magic LSB substitution method (M-LSB-SM) using multi-level encryption and achromatic component of an image," *Multimedia Tools Appl.*, vol. 75, pp. 14867–14893, 2016.

[32] K. Muhammad, J. Ahmad, H. Farman, Z. Jan, M. Sajjad, and S. W. Baik, "A secure method for color image steganography using gray-level modification and multi-level encryption," *KSII Trans. Internet Inf. Syst.*, vol. 9, pp. 1938–1962, 2015.

Khan Muhammad (S'16) received the Bachelor's degree in computer science from Islamia College Peshawar, Peshawar, Pakistan, in 2014, with a focus on information security. He is currently working toward the M.S. degree leading to Ph.D. degree in digital contents from Sejong University, Seoul, South Korea.

He has been a Research Associate with the Intelligent Media Laboratory, Seoul, South Korea, since 2015. He has authored more than 40 papers in peer-reviewed international journals and conferences, such as *Future Generation Computer Systems, Neurocomputing*, the IEEE ACCESS, the *Journal of Medical Systems, Biomedical Signal Processing and Control, Multimedia Tools and Applications, Pervasive and Mobile Computing, SpringerPlus, KSII Transactions on Internet and Information Systems*, MITA 2015, PlatCon 2016, FIT 2016, and ICNGC 2017. His research interests include image and video processing, information security, image and video steganography, video summarization, diagnostic hysteroscopy, wireless capsule endoscopy, computer vision, deep learning, and video surveillance.

Rafik Hamza received the Master's and Ph.D. degrees in cryptography and security from the Department of Computer Science, University of Batna 2, Batna, Algeria, in 2014 and 2017, respectively.

His research interests include digital images security, image and video processing, randomness algorithms, and probabilistic approaches. He has authored or coauthored several articles in these areas in reputed journals, such as the IEEE ACCESS, *Journal of Information Security and Applications*, and *Pervasive and Mobile Computing*.

Dr. Hamza is an active Reviewer of several international journals, such as Springer *Multimedia Tools and Applications*, Elsevier *Future Generation Computer Systems*, Elsevier *Computers and Electrical Engineering*, and the IEEE ACCESS.

Jamil Ahmad (S'16) received the B.C.S. degree with distinction in computer science from the University of Peshawar, Peshawar, Pakistan, in 2008, and the Master's degree with specialization in image processing from Islamia College Peshawar, Peshawar, Pakistan, in 2014. He is currently working toward the Ph.D. degree in digital contents at Sejong University, Seoul, South Korea.

He is also a regular Faculty Member with the Department of Computer Science, Islamia College Peshawar. His research interests include deep learning, medical image analysis, content-based multimedia retrieval, and computer vision. He has authored or coauthored several articles in these areas in reputed journals, including Springer *Journal of Real-Time Image Processing*, Springer *Multimedia Tools and Applications* (MTAP), Elsevier *Journal of Visual Communication and Image Representation*, *PLoS One*, Springer *Journal of Medical Systems*, Elsevier *Computers and Electrical Engineering*, *SpringerPlus*, *Journal of Sensors*, and *KSII Transactions on Internet and Information Systems*. He is also an active Reviewer for the *IET Image Processing, Engineering Applications of Artificial Intelligence, KSII Transactions on Internet and Information Systems*, MTAP, IEEE TRANSACTIONS ON IMAGE PROCESSING, and the IEEE TRANSACTIONS ON CYBERNETICS.

Jaime Lloret (M'07–SM'10) received the B.Sc.+M.Sc. degree in physics in 1997, the B.Sc.+M.Sc. degree in electronic Engineering in 2003, and the Ph.D. degree in telecommunication engineering (Dr. Ing.) in 2006.

He is currently an Associate Professor with the Polytechnic University of Valencia, Valencia, Spain. He is the Chair of the Integrated Management Coastal Research Institute and is the Head of the "Active and collaborative techniques and use of technologic resources in the education" Innovation Group. He has authored 22 book chapters and has more than 400 research papers published in national and international conferences and journals.

Dr. Lloret is an Editor-in-Chief for the *Ad Hoc and Sensor Wireless Networks* (with ISI Thomson Impact Factor) and *Networks Protocols and Algorithms*. He leaded many national and international projects. He is currently the Chair of the Working Group of the Standard IEEE 1907.1. He has been a General Chair (or Co-Chair) of 38 international workshops and conferences. He is an IARIA Fellow.

Haoxiang Wang is currently the Director and a Lead Executive Faculty Member of GoPerception Laboratory, Ithaca, NY, USA, and has been the Research Associate with Cornell University, Ithaca, NY, USA, since 2014. His research interests include multimedia information processing, pattern recognition and machine learning, remote sensing image processing, and data-driven business intelligence. He has coauthored more than 30 journal and conference papers in these fields.

Dr. Wang is a Guest Editor for several special issues.

Sung Wook Baik (M'16) received the B.S. degree in computer science from Seoul National University, Seoul, South Korea, in 1987, the M.S. degree in computer science from Northern Illinois University, Dekalb, IL, USA, in 1992, and the Ph.D. degree in information technology engineering from George Mason University, Fairfax, VA, USA, in 1999.

From 1997 to 2002, he was with Datamat Systems Research, Inc., as a Senior Scientist of the Intelligent Systems Group. In 2002, he joined the Faculty of the College of Electronics and Information Engineering, Sejong University, Seoul, South Korea, where he is currently a Full Professor and a Dean in digital contents. He is also the Head of the Intelligent Media Laboratory, Sejong University. His research interests include computer vision, multimedia, pattern recognition, machine learning, data mining, virtual reality, and computer games.