Heterogeneous and Customized Cost-Efficient Reversible Image Degradation for Green IoT

Ruoyu Zhao, Yushu Zhang, *Member, IEEE*, Rushi Lan, Zhongyun Hua, *Member, IEEE*, and Yong Xiang, *Senior Member, IEEE*

***Abstract*—With the large-scale deployment of the Internet of Things (IoT) in daily life, more and more privacy data are collected by IoT devices. These data are not directly physically controlled by users, which may cause privacy concerns. In fact, privacy has become one of the significant problems faced by IoT. In this article, we mainly study the protection of image privacy under the green IoT. We have conducted an in-depth analysis of the green IoT scenario and put forward the scope and corresponding goals that the scheme should have. Motivated by this, a novel image privacy protection scheme is proposed, i.e., heterogeneous and customized cost-efficient reversible image degradation for green IoT. This scheme fully considers the characteristics of privacy and the various users' diverse requirements to achieve a heterogeneous and customized privacy protection. Meanwhile, cost effectiveness cannot be confined to the efficiency of the direct image processing at the expense of greatly increasing costs in other aspects, such as transmission and reversion. It is mitigated by preserving some visual content in the privacy-protected image. It also improves the image compression efficiency and ensures that the user can select the desired image according to the visual content for reversion. Some experiments have been carried out to demonstrate that this work has achieved the proposed scope and corresponding goals.**

***Index Terms*—Cost efficiency, green Internet of Things (IoT), privacy protection, usability.**

Manuscript received 25 June 2022; revised 16 September 2022; accepted 9 October 2022. Date of publication 12 October 2022; date of current version 24 January 2023. This work was supported in part by the National Key Research and Development Program of China under Grant 2021YFB3100400; in part by the Research Fund of Guangxi Key Laboratory of Multi-Source Information Mining and Security under Grant MIMS20-02; in part by the Natural Science Foundation of China under Grant 62072237, Grant 62201233, and Grant 62172120; in part by the Training Program for Academic and Technical Leaders of Jiangxi Province under Grant 20204BCJL23036; and in part by the Basic Research Program of Jiangsu Province under Grant BK20201290. *(Corresponding author: Yushu Zhang.)*

Ruoyu Zhao and Yushu Zhang are with the College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China, and also with the Guangxi Key Laboratory of Multi-Source Information Mining and Security, Guangxi Normal University, Guilin 541004, China (e-mail: zhaoruoyu@nuaa.edu.cn; yushu@nuaa.edu.cn).

Rushi Lan is with the Guangxi Key Laboratory of Image and Graphic Intelligent Processing, Guilin University of Electronic Technology, Guilin 541004, China (e-mail: rslan2016@163.com).

Zhongyun Hua is with the School of Computer Science and Technology, Harbin Institute of Technology (Shenzhen), Shenzhen 518055, China (e-mail: huazhongyun@hit.edu.cn).

Yong Xiang is with the School of Information Technology, Deakin University, Burwood, VIC 3125, Australia (e-mail: yxiang@deakin.edu.au).

Digital Object Identifier 10.1109/JIOT.2022.3213875

I. Introduction

NOWADAYS, the Internet of Things (IoT) is one of the most promising and widely deployed communications networks that can collect and transmit information around them at any time [1], [2]. The concept of IoT can be traced back to ubiquitous computing, i.e., anytime and anywhere in daily life, proposed in the 1980s [3]. IoT can be applied in various services and industries, such as smart homes, automatic drives, supply chains, and smart cities. In fact, with the rapid progress of 5G and sensor technology, a large number of IoT devices, as shown in Fig. 1, have been deployed in many scenarios including the above, which may have billions of devices [4].

With the frequent exchange of information in IoT, privacy threats have become one of the biggest challenges in the development of IoT [5]. Whether individuals like it or not, a big number of IoT devices collect people's lives and pieces in the form of too detailed and finely grained images. In smart homes, for example, cameras are often placed in homes to prevent burglars or keep a watch on the elderly. These cameras inevitably capture huge portions of the owner's life and secrets; in other words, they contain much privacy. More seriously, the images collected by IoT devices are usually uploaded to third-party clouds that are not controlled by the users (i.e., privacy violators) [6] as shown in Fig. 1, and even others can access the data for some services [2].

Traditional image encryption [2], [7], [8], [9] is a method to deal with image privacy, in which the original image with abundant visual content is transformed into the encrypted one with snowflakes and like-noises. However, the basic idea of this privacy protection is that the visual content is completely abandoned, thereby causing some problems. A scenario is described in which a user installs many cameras in his or her house to monitor if the elderly fall. The cameras collect images at regular intervals (e.g., every 5 min) and transmit them to the user. It can be predicted that encrypting these images using traditional methods before the transmission will be discomfort for users. As shown in Fig. 2, no useful information can be obtained from these encrypted images before decrypting.

Intuitively, in order to comprehend the contents of the original images, users need to decrypt all of the encrypted images. However, the great majority of these gathered images may not include precise information (e.g., an image of an old person falling), making most decryption pointless. This will inevitably make the IoT systems that are already energy-intensive

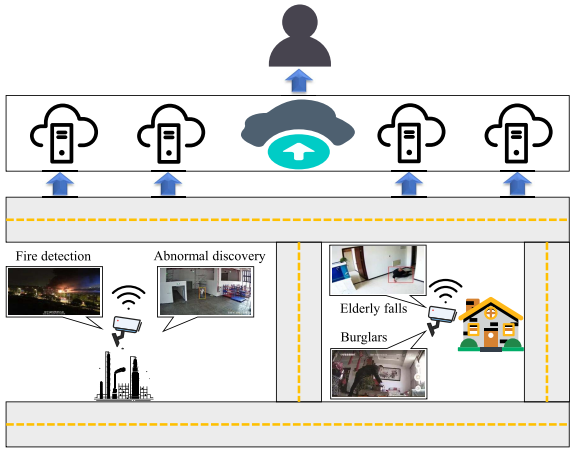


Fig. 1. Example of the IoT scenario and architecture, including data collection, cloud service, and user.

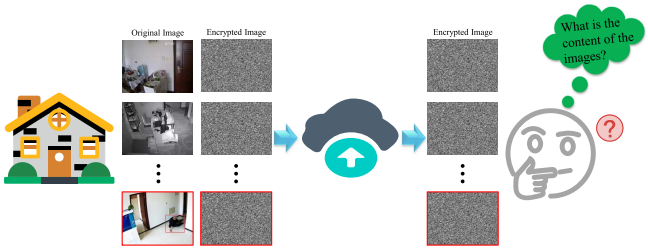


Fig. 2. Images collected by IoT are encrypted before uploading. Users cannot find the important image (i.e., red box) from the image set with the same visual effect by previewing.

consume more energy. In 2020, the IoT networks became a significant energy consumer in the information communication field. Meanwhile, it is predicted that relevant devices will consume about 45 TWh of electric energy in 2025, which is almost equal to Portugal's power consumption in 2015 [4]. Furthermore, the power consumption will continue to rise as the number of devices increases and the power consumption of a single device rises fast.

Many people agree that there is a direct relationship between energy consumption and the environment [10], mainly because fossil fuels are the primary source of energy for human beings, about 70% [11]. Meanwhile, the use of fossil fuels is a major source of carbon emissions, contributing to the greenhouse effect, severe weather, and sea-level rise. Temperatures are expected to climb by 3.4 °C by the end of the century, with carbon emissions increasing by nearly 13% between 2020 and 2030 [12]. Marine heatwaves (an extreme marine event) may become more frequent and disastrous and cause irreversible and serious changes to the ecosystem due to global warming [13].

With the awakening of humans to environmental events, IoT researchers attempt to examine ways to reduce the cost of energy consumption and protect environment in various application scenarios, termed green IoT, and the field of privacy protection is no exception. Recently, Gu et al. [2] proposed a low-power, lightweight, and precision-limited image encryption method to adapt the green IoT, namely, IEPSBP. This scheme pays attention to the cost efficiency, the low cost, and takes into account the accuracy limitation of IoT. However, it only considers the single efficiency problem in the process of privacy protection and does not consider the cost of other aspects.

As mentioned above, first, users cannot get any useful information from the images before reversing, and they have to download and decrypt a series of images to obtain one desired image. Even if a desired image is finally obtained, the bandwidth necessary to download these images and the computational cost required to reverse the image are mostly futile. That is, these expenses contribute nothing to the end results and are thus wasted. Not to mention that such tremendous efforts may not result in the desired image. Second, the image format has its own compression algorithm, which utilizes the redundancy and correlation between pixels in the image to compress the size. IEPSBP completely destroys the correlation and redundancy, making it difficult for the compression algorithm to compress the image, which leads to the rapid expansion of the image size. Compared with the original image, the size has increased by tens or even hundreds of times [14], directly increasing storage and transmission costs. As a result, while the cost of the process of privacy protection may be lowered to some extent, it raises expenses in other aspects, and the gains are not worth the loss.

To sum up, the protected image should preserve certain visual content related to the original image for the privacy protection scheme applied in green IoT. First, users can distinguish and select images of specific content to reverse according to the rough visual content by previewing, which greatly reduces the inefficient cost of transmission and reversion. Second, preserving the original visual content means that the pixels in the protected image have a certain correlation and redundancy, which makes the compression algorithm utilize this to compress the image better to reduce the image size, decreasing the inefficient cost of storage and transmission. Meanwhile, the amount of information contained in the protected image should be customized by users according to their wishes in order to better protect privacy.

Motivated by this, we propose the scheme, namely, heterogeneous and customized cost-efficient reversible image degradation for green IoT, to protect image privacy in green IoT. This scheme takes into account the varied privacy sensitivities of different components in the image, as well as the reality that different people have varying privacy viewpoints and tolerance. It allows different degrees of image degradation according to the component and people's sensitivity point of view, i.e., heterogeneous and customized. Based on this, this scheme is cost efficient, i.e., the specific image can be selected for reversion, with small expended expansion, low transmission cost, etc.

We make the following contributions in this article.

1) We deeply analyzed the scope and corresponding goals of the image privacy protection for green IoT. Cost efficiency, privacy, and usability are considered in detail. In addition, we proposed that for green IoT, the efficiency of necessary processes, such as transmission and

reversion, should be considered, which is not available in the previous related schemes.
2) We put forward a targeted image privacy protection scheme, i.e., heterogeneous and customized cost-efficient reversible image degradation for green IoT, according to the proposed scope and goals. It takes full account of the privacy of images and users' specific requirements and meets the needs of green IoT.
3) Some experiments have demonstrated that this scheme achieves the proposed scope and goals. It should be noted that it not only has cost efficiency in image degradation and reversion but also allows users to pick the desired image for reversion rather than wasting a lot of meaningless reversed costs.

We organized the rest of this article as follows. Section II briefly introduces the related works about the image degradation for privacy protection. Some basic preliminaries are provided in Section III, and meanwhile, the threat model, scope, and goal of the proposed scheme are presented in Section IV. The construction of this scheme is detailedly introduced in Section V and the evaluations are shown in Section VI. Section VII discusses the achievement of the goals of the proposed scheme and Section VIII concludes the work.

II. Related Work

Image degradation refers to the downgrading of image resolution and quality, or missing content, for various reasons. It makes the image blurred and detail difficult to perceive, thereby protecting privacy. Researchers from privacy protection communities have taken advantage of it to come up with some schemes. Here, some representative works are introduced from both irreversible and reversible aspects.

A. Irreversible Image Degradation

1) Image Filtering: This is a kind of the privacy-preserving image degradation method which is easily perceived in daily life, such as mosaic, Gaussian blur, and so on. Mosaic, also named pixelization, is a widely used privacy protection, e.g., it often appears in secret interviews to protect informants [15]. The effect of the Gaussian blurred image is similar to that of the ground glass [16] and compared with mosaic, the processed image is smoother. von Zezschwitz et al. [17] proposed a scheme to prevent bystanders from peeping into the album by distorting the images. User experiments have shown that it has good usability while safeguarding privacy, and that users can distinguish the distorted images based on the visual content.

2) Object Deletion: This method first determines (manually or automatically) the private part of the image and then erases this part [18]. Generally, some techniques (e.g., image inpainting [19]) are used to repair the erased parts to make the image representation less abrupt. Ding et al. [20] proposed a Gaussian weighted nonlocal texture similarity measurement scheme to delete the big object and then paint the image that has rich textures and geometric structures. Shetty et al. [21] considered that the existing schemes can only deal with target objects and proposed an automatic object deletion scheme that can adapt to general scenes.

3) Object Replacement: This method is to replace privacy-sensitive objects with ones related to the original. Broadly speaking, face de-identification [22] and attribute privacy [23], the hotspots of privacy protection, all belong to this category and are usually irreversible. For example, the main step of face de-identification is to generate a face with a different identity from the original image and then replace the original face. Meanwhile, some aesthetic object replacement methods are proposed. Hassan et al. [24] proposed a scheme for detecting privacy-sensitive objects and replacing them with corresponding cartoon objects that a fair compromise between deleting privacy information and keeping semantics.

4) Brief Summary: These schemes, although attempting to protect image privacy, have some problems. First, they are irreversible, which severely limits the usability of the schemes by users. For example, the images captured by the home camera may be processed using a filter method such as mosaic, uploaded to the cloud, and then broadcast to users. The user finds an image that may be abnormal but cannot understand the specific information to determine whether it is a false alarm. For object deletion, it may cause users to misinterpret the image since the content used to fill the deleted objects may be unrelated to the original content, which may fundamentally change the semantics of the visual content. Although the object replacement may mitigate this problem, it has the same fundamental problem as the object deletion, i.e., the two cannot be applied to the IoT since they require a lot of learning and computing costs.

B. Reversible Image Degradation

1) Region of Interest Encryption: This kind of method is to divide the visual content of an image into privacy-sensitive areas and public ones [25], [26]. For the privacy area, it is encrypted with content that people cannot visually grasp by previewing; for the public one, it may be published without any processing. Intuitively, users can obtain available information according to the contents of public areas, while the sensitive contents can not be understood at all, which can well protect privacy. In fact, except for the areas directly related to privacy (e.g., face), other areas carry less sensitivity, but it does not mean that there is no privacy. In other words, privacy is more like a continuous value than an absolute binary (i.e., it either has to exist or does not) [27].

2) AI-Empowered Transformation: With the proposal of reversible AI transformation methods (e.g., CycleGan [28]), privacy protection practitioners have proposed some schemes. Wu et al. [29] utilized CycleGan to transform the original visual content into the oil painting style, erasing the details and preserving the rough contents, such as outline and color. Similarly, Chai et al. [30] proposed that the image is processed by CycleGan under the condition that the original thumbnail is preserved in the processed image. While such methods protect the privacy of the visual content, as mentioned above, they cannot be applied to the IoT since AI methods require a lot of resources for computing, which is incompatible with the low energy consumption and low resources of IoT.

3) Reversible Filtering: Recently, some reversible filter methods have been proposed to protect the whole image privacy. Çiftçi et al. [31] proposed a reversible scheme by false colors to protect privacy and experiments verified the effectiveness. However, it requires a professional palette to carry out the appropriate color inpainting, which is unavailable to ordinary users. Thumbnail-preserving encryption (TPE) [32], [33], [34], [35] was proposed to protect privacy, which preserves the usability of the processed image by using format-preserving encryption to preserve the thumbnail of the image during encryption. Compared with region-of-interest (ROI) encryption, on the other hand, TPE went to the other extreme to some extent since it does not take into account the privacy sensitivity of the image contents, which are processed indiscriminately.

4) Brief Summary: Compared with Section II-A, all methods in this section are reversible, and thus they are more usability in the whole IoT scenario for users. However, they still have some problems. As mentioned above, whether there is privacy in the content is not an absolute thing but a gradual process. Obviously, the more the processing effect, the better the privacy and the worse the usability; conversely, the lesser the processing effect, the worse the privacy and the better the usability. Therefore, it should be a differentiated treatment rather than an unprocessed or indiscriminate one. In other words, each part of the image should be processed to protect privacy. However, the differential processing should be carried out according to sensitivity, i.e., heterogeneous processing, to obtain better usability.

III. Preliminary

A. TPE

TPE means preserving the original thumbnail in the processed image. First, the processes for generating thumbnails are described. Other methods of generation may not be precisely the same as the processes, but the thumbnail process generated in TPE is as follows. Note that the image generated by the following process preserves sufficient visual information no matter how the actual thumbnail is generated.

1) The image is treated as a two-dimensional (2-D) numerical matrix and divided into squares (namely, the thumbnail block) of the same size with $b \times b$.
2) The pixel values in each thumbnail block are summed and then the average is calculated.
3) The thumbnail of the image is made up of these averages.

As shown above, thumbnails may be preserved as long as the sum of the pixel values in each processed block remains constant. Intuitively, the permutation process, which changes the location of the pixels in each block, appears to be the simplest method to preserve the thumbnail. That is what the first TPE scheme did [36]. However, this operation exposes too much information about the original image in addition to the thumbnail image, such as statistics on pixel values. Some studies have suggested that the original image can be recovered from the exposed information alone [37], [38]. Therefore, it is necessary to change the value without changing the sum of the pixel values. It may seem difficult, but the sum-preserving encryption (SPE), a variant based on format-preserving encryption [39], provides a way to solve this problem. The following illustrates SPE with an example of a numeric vector with n elements ($\vec{v} = \{v_1, \ldots, v_n\}$).

1) The message space $\mathcal{M}$ of the vector $\vec{v}$ is determined. For SPE applied to TPE, $\mathcal{M} = (\mathbb{Z}_{d+1})^n$, i.e., each element value in $\vec{v}$ is not greater than d and not less than 0.
2) The sum of the element values in the vector $\vec{v}$ is calculated, that is, $s = \sum_{i=1}^{n} v_i$.
3) All vectors whose the sum of element values is s in $\mathcal{M}$ are listed to form a vector set $\Phi_{\mathcal{M}}(s)$. Thus, the problem of how to preserve the sum unchanged after processing is transformed into how to reversibly process the original vector $\vec{v}$ into a vector $\vec{c}$ in $\Phi_{\mathcal{M}}(s)$.
4) All vectors in $\Phi_{\mathcal{M}}(s)$ are ranked, i.e., each vector has a sequence number, called rank.
5) Two rank mapping functions can be constructed. One converts a vector into its corresponding rank in $\Phi_{\mathcal{M}}(s)$; and the other converts the rank into the corresponding vector, which are called $\text{rank}(\cdot)$ and $\text{rank}_s^{-1}(\cdot)$, respectively.
6) The rank r of the vector $\vec{v}$ is extracted by $\text{rank}(\cdot)$.
7) r is reversibly transformed into other legal rank values r_e, i.e., r_e can find the corresponding vector in $\Phi_{\mathcal{M}}(s)$.
8) r_e is converted into the corresponding vector $\vec{c}$ by $\text{rank}_s^{-1}(\cdot)$.

B. Chaotic System

Chaotic systems can be used to generate random numbers since they have some good features, such as unpredictability, ergodicity, and initial value sensitivity [2]. Based on the phase space, they can be classed into two classes, i.e., one-dimensional (1-D) chaotic system and multidimensional (MD) one [40]. For the former, it is simple but has weak chaotic properties; for the latter, its chaotic nature is obvious, but the structure is complex and costly. Users in the proposed scheme can choose the desired chaotic system based on their real demands, and we utilize a 2-D chaotic system as an example in this work since it achieves a compromise between cost and effect, as follows [40]:

$$\begin{cases} x_{i+1} = \cos(4 \times p_1 \times x_i \times (1 - x_i) + p_2 \times \sin(\pi \times y_i) + 1) \\ y_{i+1} = \cos(4 \times p_1 \times y_i \times (1 - y_i) + p_2 \times \sin(\pi \times x_i) + 1) \end{cases} \quad (1)$$

where p_1 and p_2 are two parameters, which, together x_0 and y_0, need to be provided by the user before running. By the way, Hua et al. [40] have carried out a large number of experiments on the system and demonstrated that it has sufficient chaotic characteristics.

IV. Threat Model, Scope, and Goal

A. Threat Model

This work mainly focuses on two threats to images gathered by IoT: one is to get useful information that users do not want others to know from images by nonlegal humans using their naked eyes; and the other is to analyze images by curious machines.

Naked Eye: After collecting images, they need to be transmitted over an unsafe channel to the cloud or client, which can be illegally intercepted by malicious people by various means to browse. Meanwhile, other people with access to cloud-stored images, e.g., administrators and image-based third-party services, may also browse images without the user's permission. They may directly extract user-related privacy information from images based on the naked eye or even widely disseminate highly privacy-related images. For example, in 2014, hackers illegally acquired pornographic unprocessed images, in which the privacy can be seen directly with the naked eye, from iCloud and quickly spread them all over the world via the Internet.

Curious Machines: When images are stored in the cloud, they may be secretly analyzed by the cloud or third-party machines, digging for information to accurately portray users, advertising, etc. However, this analysis is fast and does not require much effort on a single image. Specifically, if machines cannot analyze the contents of protected images, they will not crack maliciously. The number of images collected by IoT is so large that even if a few of them cannot be analyzed, the number of such images is huge. If the image takes much effort to crack, then the cost is huge, and there is a high probability that it will not succeed. Even if it succeeds, there may be no information in the image itself. Therefore, this is clearly not worth the loss for curious machines.

B. Scope

The scope of this work is to propose a cost-efficient green IoT image collection scheme that protects privacy while providing usability.

Cost Efficiency: This means that the cost of image processing, transmission, reversion, etc., is efficient throughout the IoT process and does not require much futile work. For example, some schemes [2] reduce the cost in the image processing phase at the expense of greatly increasing the cost in the transmission and reversion phases. Thus, the cost reduction is limited for the image in the entire IoT system.

Privacy: The difference between privacy and confidentiality is explained to help understand the privacy, although there are some overlaps. Confidentiality means that only authorized persons have access to the protected content and no other person has access to any information. Privacy is people-oriented, that is, individuals and their rights [41]. Specifically, privacy protection is not about not disclosing information to the outside world, but the amount of information disclosed should be determined by related people such as users.

Usability: It refers to the legitimate person, i.e., users, who can distinguish the protected images and detect the desired information. The usability and confidentiality are antagonistic relationships since confidentiality requires protected content not to disclose any information, resulting in being unavailable naturally.

In general, these three are mutually complementary relationships. The relationship between privacy and usability can be seen to some extent as a zero-sum game. The whole quantity of information in the image is certain, and some of it is considered to be private and protected, while the rest is revealed in the processed image, i.e., the source of usability. Meanwhile, naturally, the higher the cost effectiveness, the more effective the cost. For example, the greater the usability, the more legitimate users will be able to accurately identify the specific content in the image, and the lower expensive the useless reversion, meaning the more cost effective.

C. Goal

Based on the above scopes, specific goals for this work are set as follows.

Visual Observability: Some visual content related to the original image should be preserved in the processed image, which is the source of usability. By the way, people are more used to browsing the visual content of images than other information [32].

Heterogeneity: The degree of privacy is not the same between different contents in the image, and thus different privacy sensitivities should be treated in a heterogeneous way, that is, privacy-sensitive areas are heavier treated than less sensitive areas.

Customized: As mentioned above, the privacy is people-oriented, which varies from person to person, i.e., different people have various views, tolerance, and sensitivity [42]. It is up to the user to determine the degree of privacy protection, i.e., customized.

Balance Tunability: As indicated above, there is a zero-sum game between privacy and usability; meanwhile, everyone has different sensitivities and tolerance for this. Therefore, the balance point between the two should be determined by the user and be easy to tune.

Reversibility: The purpose of privacy protection is to prevent the disclosure of image privacy during transmission. For users, it should be expected to get the original version for the desired/interested image. Therefore, after obtaining the transmitted image, the user should have the ability to select a specific image for reversion.

Low Size: The processed image should have a low size expansion, which is conducive to less cost in transmission and storage and thus to the overall efficiency of the IoT system.

Low Time Cost: To some extent, the time cost of image processing and reversion is proportional to the complexity and energy consumption of the scheme. As a result, the image privacy protection method used in green IoT should have a low time cost.

V. Proposed Scheme Construction

A. Key Composition

In order to ensure that only authorized persons can reverse the image reversibly, a key is needed to control image degradation and reversion. For IoT devices, the computing accuracy of their terminals is often limited, such as 32 or 16 bit. Inspired by this, every input of the chaotic system should be limited, and the accuracy of each in this work is 16 bit (the same is true for computation accuracy), which is the same as in work [2]. Meanwhile, as shown in Fig. 3, there are 128 bits, which are divided into two parts, each with four small parts.

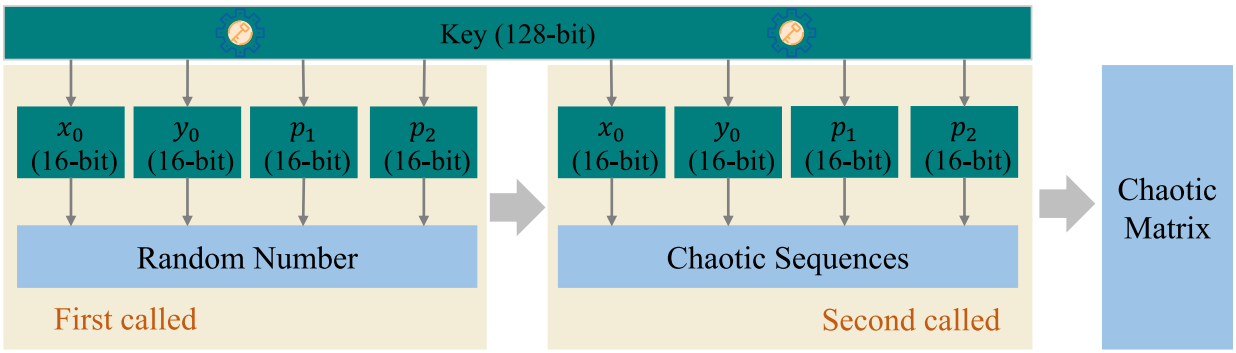


Fig. 3. Key composition and function.

Notice that the key length is not fixed and can be any length in this work. When the computation accuracy is unchanged, the length can change depending on the number of parameters needed to select the chaotic system. When the computation accuracy changes, the key length also changes. In other words, the key length varies according to the user's needs, is not fixed, and the key length is 128 bits that is just assumed in this work. Then, (1) is called twice to produce chaotic sequences and a random number, respectively. Next, how do they interact to generate the chaotic matrix is described.

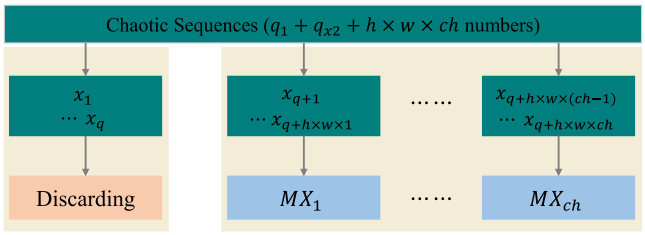


Fig. 4. Example of converting the chaotic sequences x_i into the chaotic matrix, in which $q = q_1 + q_{x2}$.

B. Chaotic Matrix

As the good characteristics of the chaotic system, such as unpredictability, ergodicity, and initial value sensitivity, the number generated by chaotic systems cannot be distinguished from true random numbers [2], making it hard for illegal third parties to reverse. Thus, the number generated by the chaotic system can be considered a random number, and the stream consisting of a series of outputs can be viewed as a chaotic sequence.

First, some researchers, on the other hand, pointed out that the output of the chaotic system has transient effects [40], i.e., the chaotic effect of the initial outputs is not obvious, as well as irregularity and instability [35]. Therefore, the first q_1 outputs of the chaotic systems need to be discarded for good effect.

Second, if the output generated by the same key only the same bit number q_1 is discarded each time, the output of the chaotic system used to process the image is the same each time. This may lead malicious people to analyze a large number of processed images and find out the pattern, and thus, the chaotic output used each time should be different. Intuitively, it can be guaranteed that the output used is different as long as the key is changed, but usually this is impossible. It means that the user processing each image needs a new key, i.e., the one-time pad, which prevents the user, not just illegal third parties, from recording so many keys. Fortunately, images have unique identifiers, e.g., shot time and filename, that can be used as the nonce t [32], i.e., the nonce t of each image is different and nonrepeating. Therefore, by calling a chaotic system for output, after discarding the first q_1 numbers, the output of the $t+1$ iterations is processed [in this work, rounding down is done after multiplying 2^{16} and calculating the absolute value as shown in (2)] to get an integer, which is called q_2. Note that (1) produces two outputs for x_i and y_i, each with its own q_2, called q_{x2} and q_{y2}, respectively

$$q_2 = \left\lfloor \mathrm{abs}\left(rn_{q_1+t+1} \times 2^{16} \right) \right\rfloor \tag{2}$$

where abs$(\cdot)$ means the absolute value, $\lfloor \cdot \rfloor$ respects the rounding down, and rn_x denotes the xth number of the output of the chaotic system.

Third, the chaotic system is again called to generate chaotic sequences, which should discard the first $q = q_1 + q_2$ outputs, Meanwhile, the number of system iterations is image dependent, e.g., the image is size $h \times w$ and has ch channels, then sequences x and y should form ch matrices, i.e., XM_i and YM_i $(i = 1 \cdots \mathrm{ch})$, respectively, in which each matrix has $h \times w$ elements. In other words, for the sequences composed of x_i, the elements are $\{x_{q+1}, \ldots, x_{q+h\times w\times \mathrm{ch}}\}$; for the sequences composed of y_i, the elements are $\{y_{q+1}, \ldots, y_{q+h\times w\times \mathrm{ch}}\}$. Then, each sequence is divided into three equal parts and each part constitutes a matrix with size $h \times w$ as shown in Fig. 4 (as an example of the sequences x_i).

C. Image Degradation

The overview of the proposed scheme is shown in Fig. 5. First, the sensitive area of the image is manually or automatically selected and determined. Of course, even if it is automatically determined via some algorithms, which also should require the user to tell the algorithm in advance what should be determined, since the algorithm is only an objective technology and it cannot know the sensitive points in the user's mind. Second, color images are often composed of several channels (usually three) and each channel can be processed

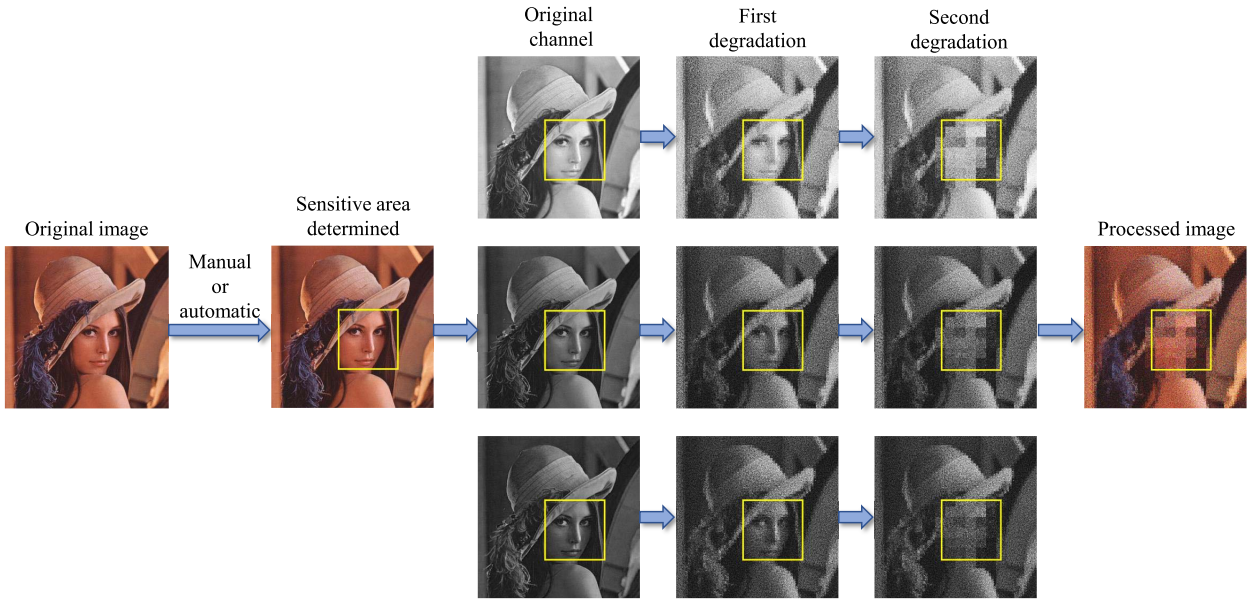


Fig. 5. Overview of the proposed image degradation.

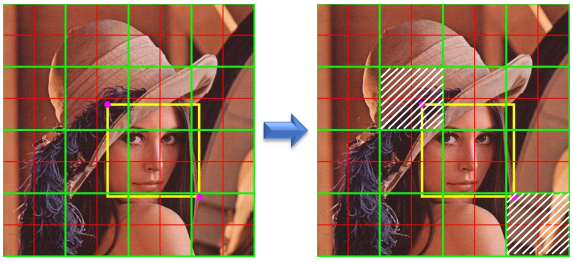


Fig. 6. Example of the determination of saved coordinates, in which the green frame means ϱ_2, the red frame denotes ϱ_1, pink dots represent the upper left and lower right corners of the sensitive area, and the white underline shows the area whose coordinates are saved in ϱ_0.

in the same way, and thus the following is the processing method of one channel unless otherwise specified. Third, the first degradation is performed on the whole image, and subsequently, the second degradation is performed on the selected area.

It can be seen from Fig. 5 that the degradation of this work is carried out in blocks. Specifically, the correlation between elements in the block is disrupted and the correlation between blocks is preserved, i.e., the fine visual content within blocks is eliminated and the coarse content between blocks is preserved. In this work, the channel of the image is divided into three different types of blocks, ϱ_0, ϱ_1, and ϱ_2. After demarcating the sensitive area, the required coordinates need to be saved in ϱ_0 to determine the location when reversing. ϱ_1 is treated at the first degradation and ϱ_2 is treated at the second one. The following is a detailed description, and for simplicity, the block size of ϱ_2 with $b_2 \times b_2$ is a multiple of ϱ_1 with $b_1 \times b_1$.

The image is divided into blocks with the size of $b_2 \times b_2$ after determining the sensitive area. Subsequently, the coordinates of the two blocks ϱ_2 to be saved in ϱ_0 are determined according to the upper left κ_l and lower right κ_r corners of the sensitive area as shown in Fig. 6. To put it simply, after the image is divided into blocks with $b_2 \times b_2$, the coordinates of the block where the upper left corner and the lower right corner (the pink dot in Fig. 6) of the sensitive area are located are the coordinates of the sensitive area to be recorded. In a 2-D matrix, each coordinate consists of two elements, i.e., (τ_x, τ_y). The two coordinates of ϱ_2 corresponding to κ_l and κ_r are saved, called (τ_{xl}, τ_{yl}) and (τ_{xr}, τ_{yr}), respectively.

The coordinates are encoded in binary form and stored in the least significant bits (LSBs) in ϱ_1, which block(s) is/are called ϱ_0. The total length l of bits required is as follows:

$$\begin{aligned} l &= \left\lceil \log_2\left(\frac{b_2}{h}\right)\right\rceil + \left\lceil \log_2\left(\frac{b_2}{w}\right)\right\rceil \\ &\quad + \left\lceil \log_2\left(\frac{b_2}{h}\right)\right\rceil + \left\lceil \log_2\left(\frac{b_2}{w}\right)\right\rceil \\ &= \left(\left\lceil \log_2\left(\frac{b_2}{w}\right)\right\rceil + \left\lceil \log_2\left(\frac{b_2}{h}\right)\right\rceil\right) \times 2 \end{aligned} \tag{3}$$

where the four parts in the first line represent the length of bits required for τ_{xl}, τ_{yl}, τ_{xr}, and τ_{yr}, respectively. Obviously, the lengths of τ_{xl} and τ_{xr}, and τ_{yl} and τ_{yr} are equal, and thus they can be abbreviated. By the way, $\lceil\cdot\rceil$ means the rounding up to an integer. According to l, the quantity l_{ϱ_0} of ϱ_0 can be determined as follows:

$$l_{\varrho_0} = \left\lceil \frac{l}{b_1 \times b_1} \right\rceil. \tag{4}$$

Normally, $l_{\varrho_0} = 1$, since it often means that the size ($b_1 \times b_1$ and $b_2 \times b_2$) of the block is too small compared to the size ($h \times w$) of the image if it is greater than 1. Meanwhile, it is worth declaring that this work also allows the existence of $l_{\varrho_0} > 1$ without change, but this situation is relatively rare in practice. For images with multiple channels, since the positions of sensitive areas are the same, it is only necessary to store the positions on one channel, that is, ϱ_0 only exists on one and not the other channels. Meanwhile, for the convenience of description, it is assumed that ϱ_0 does not coincide

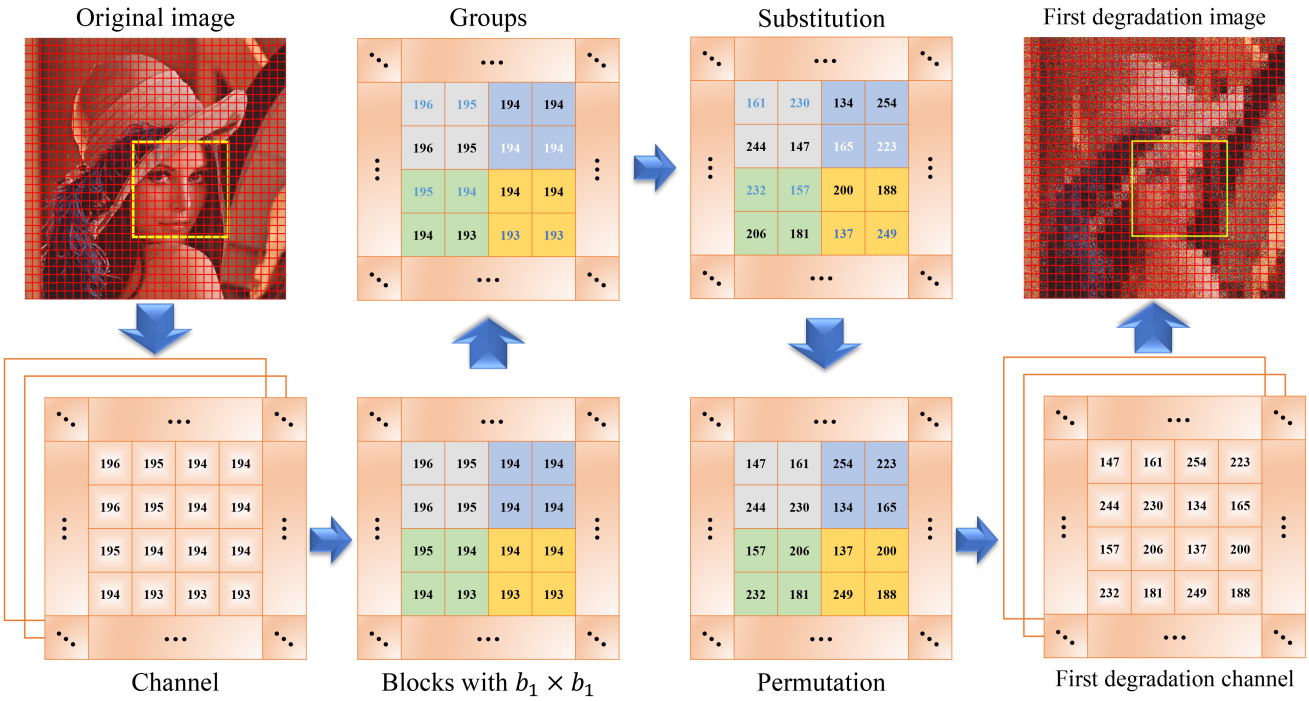


Fig. 7. Overview of the first degradation.

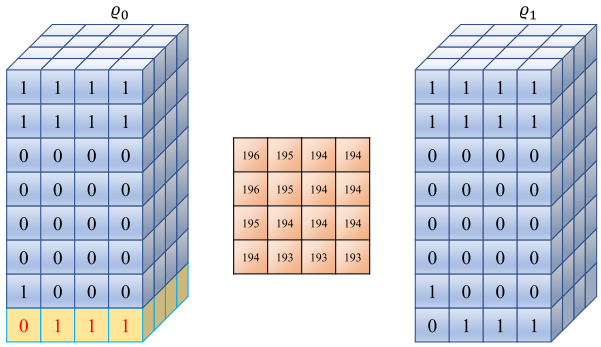


Fig. 8. Difference in the number of processing bits in ϱ_0 and ϱ_1, where the blue cube represents the bits to be processed in substitution and permutation.

with ϱ_2, which is consistent with most actual situations. In fact, even if it coincides, the processing method is similar.

Then, the first degradation in ϱ_0 and ϱ_1 is described in Fig. 7 and it can be divided into two main steps: 1) substitution and 2) permutation. Note that in the two steps as shown in Fig. 8, for ϱ_0, the upper 7 bits are processed, and for ϱ_1, the entire element (i.e., 8 bits) is processed. For the substitution, the SPE introduced in Section III-A is utilized. Although it appears that all elements in each block can be directly processed at one time, this results in too many vectors in the set, thereby too high computational cost, which rises rapidly with the increase of the block size. Fortunately, the computational cost of processing two elements at once is very low, making it applicable to IoT. Based on the group within two elements, i.e., (ι_a, ι_b), rank mapping is designed as follows:

$$\text{rank}(\iota_a, \iota_b) = \begin{cases} \iota_a, & s \le d \\ \iota_a - s + d, & \text{otherwise} \end{cases} \tag{5}$$

$$\text{rank}_s^{-1}(r) = \begin{cases} (r, s - r), & s \le d \\ (s - d + r, d - r), & \text{otherwise} \end{cases} \tag{6}$$

where $s = \iota_a + \iota_b$, $d = 127$ for ϱ_0, $d = 255$ for ϱ_1. With (5) and (6), the group and its rank can be converted to each other. After extracting its rank r_o from (ι_a, ι_b), r_o is encrypted as follows:

$$r_e = \text{mod}(r_o + \varepsilon, |\Phi_{\mathcal{M}}(s)|) \tag{7}$$

where $\text{mod}(\cdot, \cdot)$ means the remainder function, $|\cdot|$ represents the number of vectors in the set, and ε denotes an enough large number. For a vector/group with two elements, its $|\Phi_{\mathcal{M}}(s)|$ is as follows:

$$|\Phi_{\mathcal{M}}(s)| = \begin{cases} s + 1, & s \le d \\ 2 \times d - s + 1, & \text{otherwise.} \end{cases} \tag{8}$$

For ε, the chaotic matrix XM_i is used. Since XM_i and the channel are 2-D matrices of the same size, each group (ι_a, ι_b) on the channel corresponds to a group (ι_{xa}, ι_{xb}) on XM_i, and ε is calculated as follows:

$$\varepsilon = \lfloor (\text{abs}(\iota_{xa}) + \text{abs}(\iota_{xb})) \times \mu \rfloor \tag{9}$$

where μ means an enough large number selected by users. The substitution operation is complete when this is done for each group in ϱ_0 and ϱ_1.

The operation of permutation is illustrated in Fig. 9, showing that for ϱ_0, as with substitution, only the higher seven bits are processed, while for ϱ_1, the entire element is processed. In this operation, the chaotic matrix YM_i is used, which is divided into blocks with size $b_1 \times b_1$, and then the elements in the block are ordered by size. Last, the elements in ϱ_0 and ϱ_1 that participate in the permutation change their positions according to this order. After the above operations, the first degradation is completed.

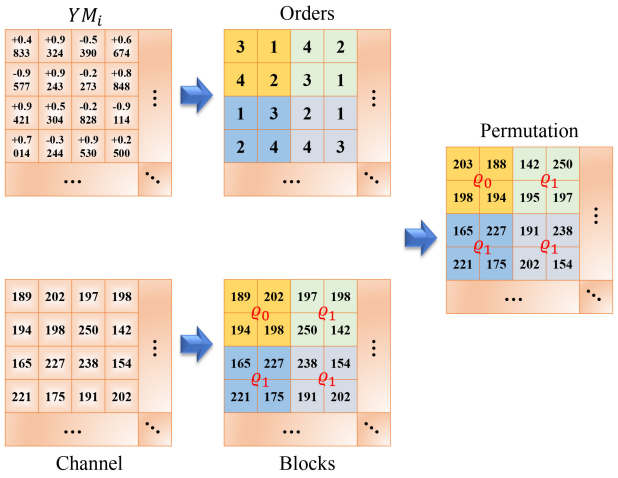


Fig. 9. Example of the permutation in ϱ_0 and ϱ_1.

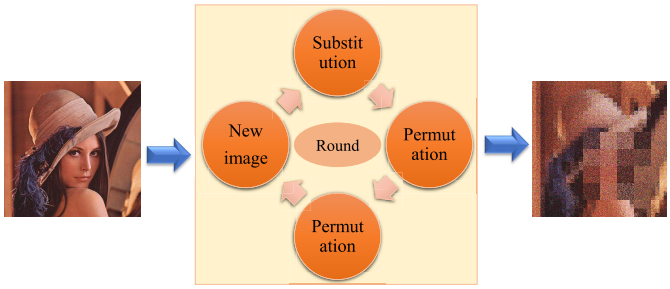


Fig. 10. Example of the image degradation round.

The second degradation is for the sensitive area. After dividing the area into blocks with size $b_2 \times b_2$, the same permutation operation as above is carried out in combination with YM_i. The permutation operation in Fig. 7 can also illustrate the second degradation. The first and the second degradation together, i.e., a substitution and two permutations, are called a degradation round as shown in Fig. 10. Users can carry out multiple rounds of the image degradation in accordance with their actual needs.

If the same chaotic matrix is used for the two rounds, the attacker may analyze it and get some rules. To prevent this potential risk, there are two simple methods that can be utilized when performing multiple rounds of degradation as follows.

1) A sequence number is added to the input of a chaotic system, which changes with the number of rounds. For example, for each image degradation round, the sequence number is increased by 1. Therefore, for the same image, the input to the chaotic system is different in each degradation, and thus the output chaotic sequence is also different. By the way, some schemes [32], [33], [43], [44] that require multiple rounds of operation are also used in this way to ensure that the chaotic sequence used is different each time.
2) The number of the round, t_r, to degrade is preset and a chaotic sequence that the length is $(q_1 + q_{x2} + h \times w \times \text{ch}) \times t_r$ is generated. Then, the chaotic sequence is averaged into t_r segments which are processed as shown in Fig. 4.

Since a chaotic system is sensitive to initial values, unpredictable, and random-like, there two methods can ensure that the chaotic matrices used in each degradation round are different and irregular.

Subsequently, the LSBs of ϱ_0 are processed, and notice that the coordinates of the sensitive area, i.e., (τ_{xl}, τ_{yl}) and (τ_{xr}, τ_{yr}), have been converted into binary bits and stored in the LSBs. The areas with size $b_1 \times b_1$ corresponding to ϱ_0 from the corresponding XM_i and YM_i of the channel where ϱ_0 is located are called XM_{ϱ_0} and YM_{ϱ_0}, respectively. They are treated as follows:

$$M_{\text{xor}} = \begin{cases} 0, \ \text{XM}_{\varrho_0}(x, y) + \text{YM}_{\varrho_0}(x, y) \leq 0 \\ 1, \ \text{otherwise} \end{cases} \tag{10}$$

where $M(x, y)$ means the element of row x and column y in 2-D matrix M. Then, the LSBs of ϱ_0 and M_{xor} are XORed as follows:

$$\text{LSB}_e(x, y) = \text{LSB}_o(x, y) \oplus M_{\text{xor}}(x, y), (1 \leq x, y \leq b_1) \tag{11}$$

where LSB_o means the LSBs of ϱ_0 and $\oplus$ denotes the operator of XOR.

D. Reversion

In general, the reverse process of image degradation is the same as the above steps, but in reverse orders as follows.

1) The LSBs of ϱ_0 can be reversed by XOR operation again after the chaotic matrix is generated.
2) (τ_{xl}, τ_{yl}) and $(\tau_{xr}, \tau_{\text{yr}})$ are extracted from the LSBs, i.e., the sensitive area can be determined.
3) The sensitive area is divided into blocks ϱ_2 with $b_2 \times b_2$ and then the inverse permutation is performed in ϱ_2.
4) The channel is divided into blocks ϱ_1 with $b_1 \times b_1$ and then the inverse permutation is performed in ϱ_1.
5) The inverse of the substitution is performed in groups with two elements. Specifically, the group's rank r_e is obtained by $\text{rank}(\cdot, \cdot)$ and then it is reversed as follows:

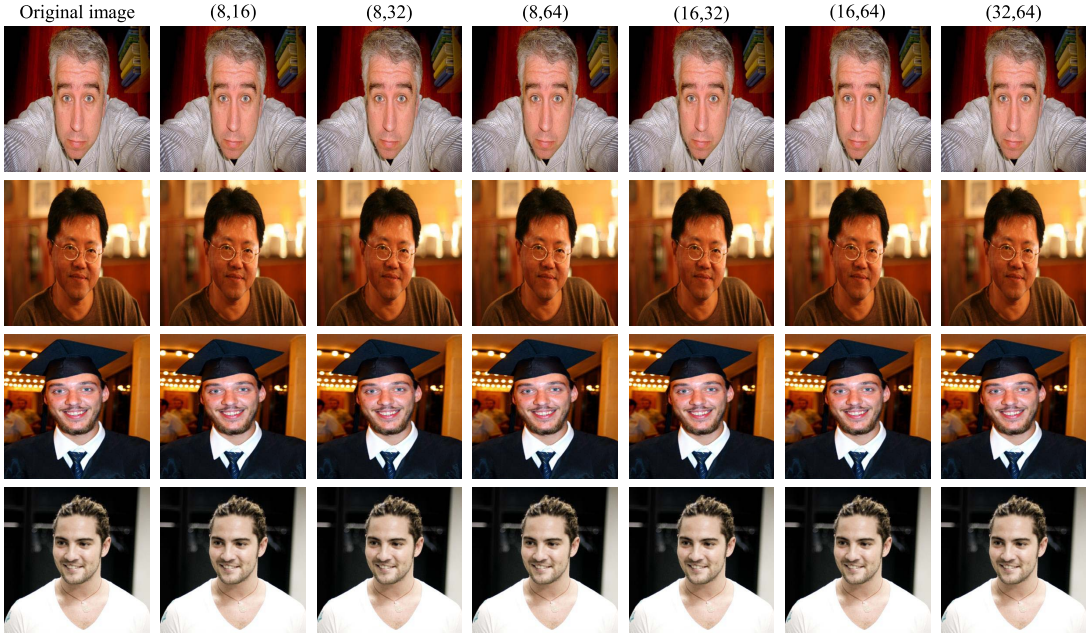
$$r_o = \text{mod}(r_e - \varepsilon, |\Phi_{\mathcal{M}}(s)|). \tag{12}$$

6) $\text{rank}_s^{-1}(\cdot)$ is utilized to convert r_o to the corresponding element groups.
7) The reversion process is completed after all groups are performed the inverse of the substitution.

VI. Scheme Evaluation

In this section, experiments and analysis are utilized to show and evaluate the excellence of this scheme. The experiments are carried out on the MATLAB 2021b with i7-8700 CPU @ 3.2 GHz, 16-GB RAM, and Window 10 platform. In experiments, the face is used as a sensitive area, and the first 1000 images in the Helen data set [45] are applied as experimental images with the PNG format, which are resized to 512×512. By the way, the illustrations in this section are also from this data set. (b_1, b_2) indicates that the block size of the first degradation is b_1 and the second is b_2, and six different combinations are tested in the experiments, i.e., $(8, 16)$, $(8, 32)$, $(8, 64)$, $(16, 32)$, $(16, 64)$, and $(32, 64)$.

(a)



(b)

Fig. 11. Visual effect of the proposed scheme: (a) degraded images and (b) reversed images.

A. *Visual Quality*

The degraded and the corresponding reversed images with different block sizes are shown in Fig. 11. Visually, the amount of information in degraded images leaked can be controlled based on changes in block sizes, while the sensitive area can be further degraded to reveal less information than others. Meanwhile, there is no difference in the visual effect between the reversed image and the original one regardless of the block size.

Two commonly used image quality evaluation indicators, structural similarity (SSIM) and peak signal-to-noise ratio (PSNR), are applied to assess the quality of reversed images, and the results are shown in Fig. 12. For SSIM, the values in Fig. 12(a) are infinitely close to 1, which means the image quality is very high as the maximum value of SSIM is 1. For PSNR, a value greater than 40 means that the image quality is very high and the difference is nearly imperceptible to the naked eye [46]; thus, the values in Fig. 12(b) denote that the reversed image quality is excellent. In all, as seen in Fig. 12, the larger the b_2, the higher the index for the reversed image; the smaller, the lower the index, which is since the larger the b_2, the fewer bits required to store in the LSBs of

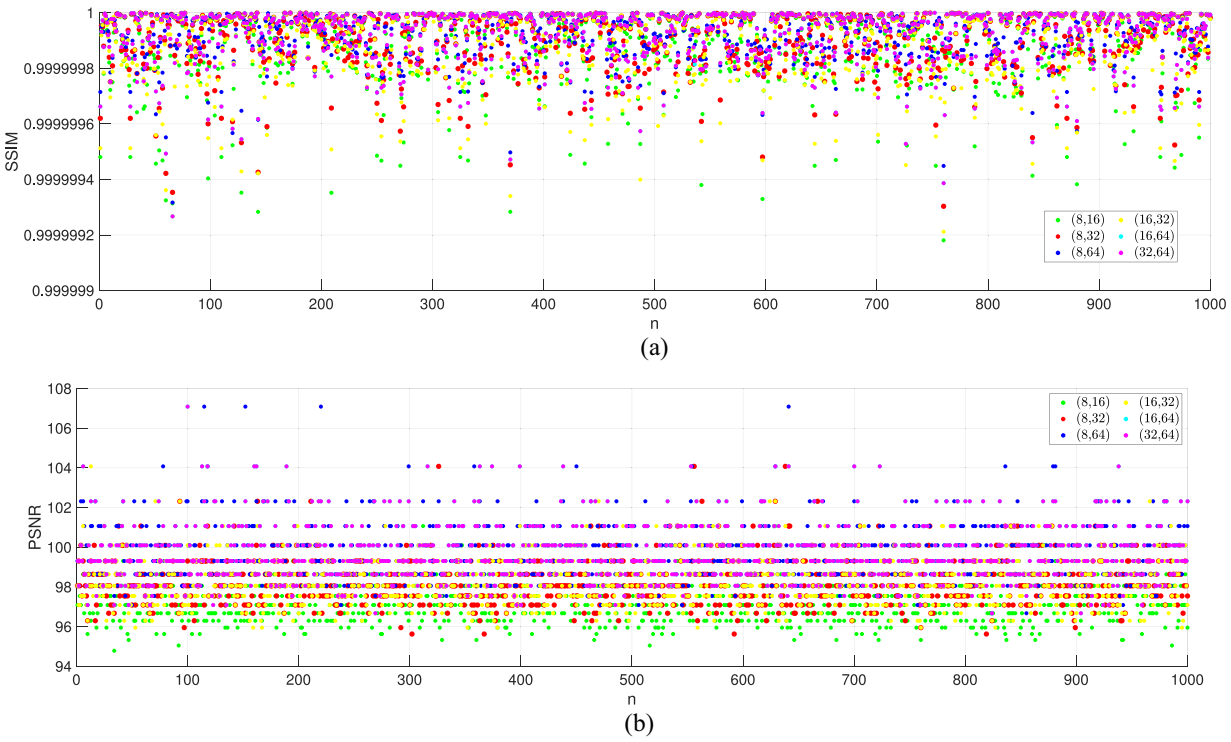


Fig. 12. Quality indicators for the reversed images: (a) SSIM and (b) PSNR.

TABLE I
TIME STATISTICS OF THE PROPOSED SCHEME

Scheme / Item	The proposed scheme					
	$(8,16)$	$(8,32)$	$(8,64)$	$(16,32)$	$(16,64)$	$(32,64)$
Degraded average time (s)	0.9051	0.9141	0.9152	0.8166	0.8107	0.7617
Reversed average time (s)	1.2130	1.2951	1.2986	1.3947	1.4898	2.5053
Degraded variance time (s^2)	0.0008	0.0012	0.0015	0.0015	0.0015	0.0003
Reversed variance time (s^2)	0.0045	0.0302	0.0742	0.0241	0.1534	0.3891
Degraded maximum time (s)	1.0460	1.0353	1.0775	0.9799	0.9954	0.8898
Reversed maximum time (s)	1.8080	2.0767	3.8895	2.2846	4.0116	5.7848
Degraded minimum time (s)	0.8648	0.8615	0.8639	0.7604	0.7600	0.7286
Reversed minimum time (s)	1.1377	1.1418	1.1445	1.2550	1.2523	2.0864
Degraded average speed $(Mbps)$	2.2097	2.1879	2.1853	2.4491	2.4669	2.6257
Reversed average speed $(Mbps)$	1.6489	1.5443	1.5401	1.4340	1.3425	0.7983

ϱ_0 as shown in (3). This is particularly evident in the indicator of PSNR as shown in Fig. 12(b), which shows that the experimental results have obvious stratification. That is, the lower PSNR value is often the case of smaller block size [e.g., (8, 16)]; the higher PSNR value is often the case of bigger block size [e.g., (32, 64)].

B. Time Cost

In this work, the time cost is guaranteed in two ways. First, all computational accuracy is limited to 16 bit, which greatly reduces unnecessary high-precision computations and enables the proposed scheme to be used on IoT devices with few resources. Second, the degraded images are visually observable, which means that users can reverse the images that are precisely selected by browsing, thereby greatly reducing futile reversions. The time cost of this work is further illustrated by the experiments below.

The time statistics on degradation and reversion are carried out, including average, variance, maximum, minimum, and average speed, as shown in Table I. It can be seen that it takes less than 1 *s*, or (*s*), for an image to degrade; most of it takes slightly more than 1 (*s*) to reverse. Although the reversed time is slighter than the degradation time, in practice, users do not want to reverse all the images as mentioned above, only a part or even a tiny part of them, and thus the reversed efficiency is considerable. The time variance, maximum, and minimum values show that the time consumed by this scheme is stable,

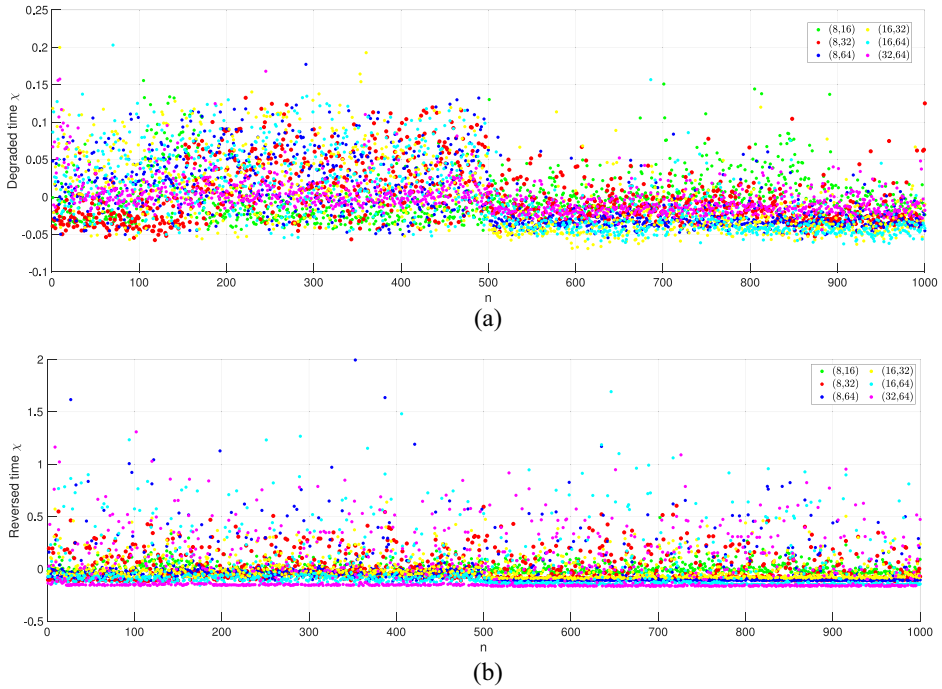


Fig. 13. Results of time fluctuations χ: (a) degraded and (b) reversed.

TABLE II
FILE SIZE OF THE PROPOSED SCHEME

Scheme / Item	The proposed scheme						
	original	$(8,16)$	$(8,32)$	$(8,64)$	$(16,32)$	$(16,64)$	$(32,64)$
Degraded size (MB)	315	670	672	675	676	679	684
Reversed size (MB)	315	315	315	315	315	315	315

especially for degraded operations. Meanwhile, the degradation average speed is about 2.2–2.6 (*Mb/s*) under different parameters. Compared with the other scheme [47] applied in IoT for image privacy protection, in which the protection speed is about 1.31 (*Mb/s*), the speed of this work is excellent. Although the reversed speed is not that high, the user actually reverses only a small portion, which offsets the difference in speed.

In addition, to further evaluate the time stability, the fluctuation rate χ of each image processing time is calculated as follows:

$$\chi = \frac{\zeta_i - \zeta_m}{\zeta_m} \tag{13}$$

where ζ_i means the time spent processing image i and ζ_m denotes the average under the same parameter. Fig. 13 shows time fluctuations under degradation and reversion operations. For the degradation as shown in Fig. 13(a), they fall within the range of ± 0.05 in most cases, denoting that degradation time is very stable; for the reversion, as shown in Fig. 13(b), although a few fluctuations are more pronounced (in fact, in which most of them are below 0.9), most of them are concentrated around 0. The experiment demonstrates that the image degraded speed is fast and stable, and the reversed speed matches it since users can select part images instead of reversing them all.

C. File Size

In this work, the file size is low in two ways. First, although the pixel correlation within a block in a degraded image is broken to protect privacy, there is still a coarse correlation between blocks. Compression algorithms within the image format can compress file sizes based on these correlations, and thus the size expansion of the degraded image remains at a low level. Second, the reversed image is almost identical to the original one, except that the LSBs of several pixels may change [most l bits as shown in (3)], and therefore, the reversed image and the original one should be almost identical in size. In addition, for degraded images, the larger the block size, the larger the file size, since this results in a reduction in the number of blocks, which reduces the correlation between blocks, thereby reducing in compression efficiency.

The results of the file size experiment on the data set are shown in Table II. For degraded images, it can be seen that as the block size increases, the file size increases, but very slowly. Meanwhile, compared with the original images, the size of the degraded image is only about doubled, far less than the image privacy protection scheme proposed by Gu et al. [2]. For reversed images, although there is a loss compared with the original image, this loss can be ignored, and thus it is not reflected in the file size, which indirectly demonstrates the high visual quality of the reversed image.

D. Key Evaluation

The key evaluation should be considered in two aspects. First, the key space should be large enough, i.e., the key is long enough, to prevent brute force cracking. Second, the key sensitivity should be strong enough that a slight change in the key can result in a significant change in either the degradation or reversion.

For the key space, the chaotic system is called twice to generate the chaotic matrix needed for degradation. As shown in Fig. 3, one 64-bit key is applied in one call and thus the key with a total of 128 bits is used for one-time image degradation. That is, the key space of the proposed scheme is 2^{128}, which exceeds the security requirement of the key space 2^{100} [2]. Thus, this scheme meets the key requirements for preventing brute-force attacks.

For the key sensitivity, the chaotic system [40] used in this work has been tested extensively, demonstrating to have the characteristics of initial value sensitivity, chaos, and unpredictability. Therefore, the chaotic sequences output by the chaotic system will be completely different as long as the key changes a little, resulting in major changes in the values in the chaotic matrix and then great changes in the output image.

The experiment on key sensitivity is carried out to illustrate this proper as shown in Fig. 14. First, two 128-bit keys, only one bit of which is different, are set as follows:

$$K_1 = 0 \times FFF\mathbf{F}FFFFFFFFFFFFFFFFFFFFFFFFFFFF$$
$$K_2 = 0 \times FFF\mathbf{E}FFFFFFFFFFFFFFFFFFFFFFFFFFFF.$$

They are used to degrade the same image and reverse the two degraded images, respectively, as shown in Fig. 14(a), denoting that the reversion process is completely ineffective even if the key is only a little different. Meanwhile, the result obtained by subtracting two degraded images is shown in Fig. 14(b), illustrating that even if the keys differ a little, the degraded images are very different. The histogram of the result also shows that a large number of pixel values have changed as exhibited in Fig. 14(c). Therefore, experiments show that the proposed scheme has the sufficient key sensitivity.

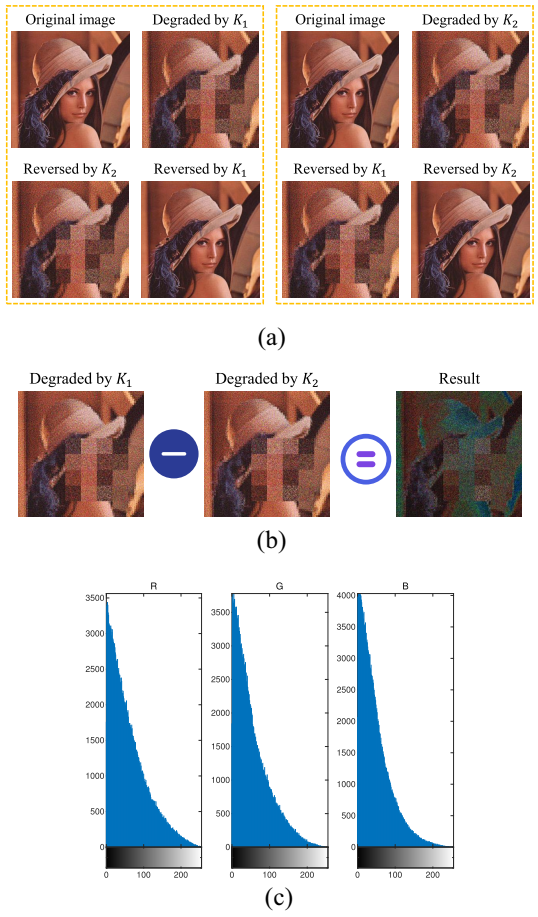


Fig. 14. Key sensitivity set: (a) example of degradation and reversion by K_1 and K_2; (b) result of subtracting two images degraded by K_1 and K_2; and (c) histogram of three channels of the result.

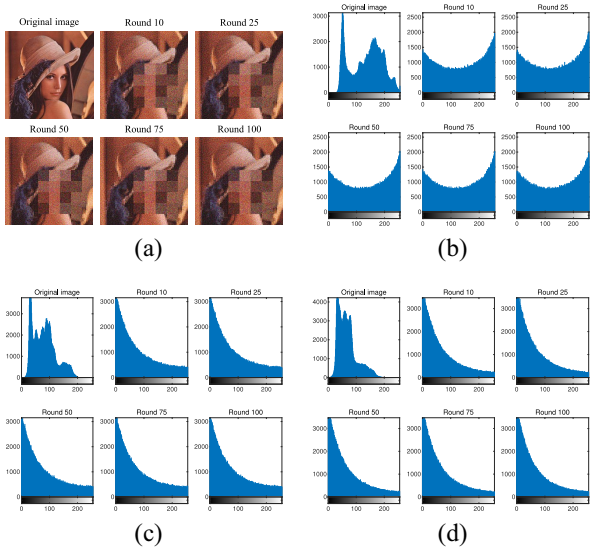


Fig. 15. Images and the histograms of their channels: (a) images; (b) R channel; (c) G channel; and (c) B channel.

E. Histogram Analysis

The proposed scheme can carry out multiple rounds of degradation according to the needs of users. In order to prevent third parties from distinguishing the number of rounds of image degradation based on the statistical information, it should not significantly differ from each round and should obviously differ from the original image. The pixel histogram is an important form and tool in statistics. As shown in Fig. 15, the histogram of each channel of the original image is irregular and fluctuates obviously. For the degraded image, on the other hand, even after different rounds, the histograms in the same channel are almost indistinguishable and very stable. In addition, the change intensity φ of pixel value frequency under different rounds is computed as follows:

$$\varphi = \frac{1}{n} \times \sum_{i=1}^{n} \left(\phi_{\text{round}_x}(i) - \phi_{\text{round}_y}(i)\right)^2 \tag{14}$$

where ϕ_{round_x} means the frequency table of the pixel value of the image degraded by x round, and for an 8-bit-per-pixel image, n is 256. For an image, the intensity change can be understood as the difference of its pixel value histogram. The

TABLE III
CHANGE INTENSITY OF PIXEL VALUE FREQUENCY IN B CHANNEL

round	0	10	25	50	75	100
0	0	1.73×10^6	1.75×10^6	1.74×10^6	1.74×10^6	1.75×10^6
10	1.73×10^6	0	2.39×2^3	2.24×2^3	2.04×2^3	2.03×2^3
25	1.75×10^6	2.39×2^3	0	2.36×2^3	2.44×2^3	1.82×2^3
50	1.74×10^6	2.24×2^3	2.36×2^3	0	2.09×2^3	1.70×2^3
75	1.74×10^6	2.04×2^3	2.44×2^3	2.09×2^3	0	1.67×2^3
100	1.75×10^6	2.03×2^3	1.82×2^3	1.70×2^3	1.67×2^3	0

TABLE IV
NUMBER OF THE FACE DETECTION FAILURE OF THE PROPOSED SCHEME

Scheme / Item	The proposed scheme						
	original	$(8, 16)$	$(8, 32)$	$(8, 64)$	$(16, 32)$	$(16, 64)$	$(32, 64)$
Numbers	0	371	760	967	886	981	999

B channels of images in Fig. 15(a) are done to the φ computation and the results are shown in Table III. The change intensity of the pixel frequency is high compared with the original image and the image after any round of degradation. In comparison between degraded images, the change intensity of the pixel frequency is stable and small. It, together with the histogram, shows that the pixel statistics of each round of images are similar but not exactly the same. This not only prevents the third party from distinguishing the number of rounds from the statistical information but also ensures changes in each round.

F. Face Detection

In Section IV-A, the curious machine is stated as a model that threatens image privacy. Specifically, the image may be simply and quickly scanned by the machine due to curiosity, but it will not be analyzed in detail. In this part, a face detection model is used as the curious machine since 1) faces are highly related to privacy, which is consistent with the purpose of protecting privacy for this work and 2) the face detection model is relatively mature compared to other models due to its practicality.

One face detection API, face++,[1] is used for this work. Since there may be more than one face in an image, for statistical purposes, we tested all images with only one sensitive are. In other words, there should be 1000 sensitive areas in the original image. The higher the number of the detection failure of these sensitive areas in the degraded image, the better the protection effect will be. The experimental results are shown in Table IV, illustrating that the larger the block size, the higher the success rate against the curious machine. This also validates the user's ability to change the intensity of privacy protection by resizing blocks.

VII. DISCUSSION

In Section IV, some goals have been proposed to achieve the scope of the scheme, i.e., cost efficiency, privacy, and usability. Here, the implementations of the proposed goals are analyzed.

Visual Observability: The proposed work first divides the image into blocks, which eliminates the fine visual information within the blocks and preserves the rough (holistic) visual information between the blocks. As shown in Fig. 11(a), the degraded image has rough visual information related to the original image, which enables users to observe through vision. That is, the goal of the visual observability is achieved.

Heterogeneity: In this work, the image can be divided into different parts, and each part is degraded according to its privacy sensitivity, i.e., the heterogeneous image degradation. Therefore, the goal of the heterogeneity is achieved.

Customized: In the proposed scheme, the selection of the sensitive area, the determination of the sensitive information, and the degree of degrading are all decided by the user. In other words, users can customize visual observability, heterogeneity, privacy, and usability according to their own wishes. In particular, it ensures that privacy has been protected as users have customized the protection degree of privacy that is subjective. Thus, the goal of the customized is implemented.

Balance Tunability: Both privacy and usability are based on visual observability in degraded images. The degradation effect of the image can be controlled by changing the block size. As shown in Fig. 16, the exposure degree of information in the degraded image can be changed by changing the block size. Thus, the balance between privacy and usability can be tuned easily.

Reversibility: In this work, the degraded image can be reversed after the user selects a specific image according to the visual effect. As shown in Fig. 11(b), the reversed image has no difference in visual effect from the original image. Meanwhile, the experiments of two image quality indicators also denote that the reserved image has the good quality as shown in Fig. 12. As a result, the goal of reversibility is realized.

Low Size: Although the correlation within the block is erased in the degraded image, the correlation between blocks is preserved. Thus, the compression algorithm within the format can utilize this correlation to compress the image size, making the degraded image at a low size level. Meanwhile, the reversed image is almost the same as the original one and thus

[1] https://www.faceplusplus.com.cn/

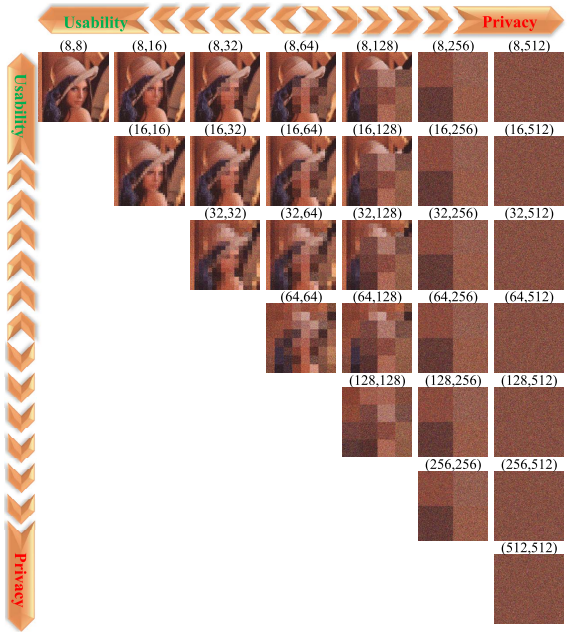


Fig. 16. Balance between privacy and usability.

the two should be almost the same size. Experimental results also demonstrate it as shown in Table II. In other words, the goal of the low size is implemented.

Low Time Cost: The chaotic system utilized in this work has the characteristics of efficiency, and the operation of the substitution and permutation is computing efficiency, and thus time cost is low. The experiment on the time cost also demonstrates it as shown in Section VI-B. Therefore, the goal of the low time cost is achieved.

VIII. Conclusion

As IoT devices become more widely and thoroughly integrated into daily life, the data they gather increasingly threatens personal privacy. Data privacy has emerged as one of the most pressing challenges confronting IoT. Image privacy in IoT has been considered and protected in this work. The target solution is necessary since the IoT terminal has limited resources and computation accuracy.

In this article, we propose a novel image privacy protection scheme for green IoT, i.e., heterogeneous and customized cost-efficient reversible image degradation. First, the privacy-protected image preserves the visual information related to the original image to ensure that users can browse. Second, the calculation cost of this scheme is efficient and friendly. In particular, it should be pointed out that this scheme can carry out targeted image reversion according to the user's ideas, thus avoiding a large number of futile reversion costs. Third, the different privacy sensitivities of different parts of the image and the different privacy needs of users are considered, and inspired by this, the new scheme allows the heterogeneous and customized image degradation. The experiments and analysis show that this work has achieved the proposed scope of cost efficiency, privacy, and usability, and meanwhile, the corresponding goals have been well achieved.

References

[1] A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi, "Internet of Things for smart cities," *IEEE Internet Things J.*, vol. 1, no. 1, pp. 22–32, Feb. 2014.

[2] Z. Gu et al., "IEPSBP: A cost-efficient image encryption algorithm based on parallel chaotic system for green IoT," *IEEE Trans. Green Commun. Netw.*, vol. 6, no. 1, pp. 89–106, Mar. 2022.

[3] L. Atzori, A. Iera, and G. Morabito, "The Internet of Things: A survey," *Comput. Netw.*, vol. 54, no. 15, pp. 2787–2805, 2010.

[4] X. Liu and N. Ansari, "Toward green IoT: Energy solutions and key challenges," *IEEE Commun. Mag.*, vol. 57, no. 3, pp. 104–110, Mar. 2019.

[5] C. Li and B. Palanisamy, "Privacy in Internet of Things: From principles to technologies," *IEEE Internet Things J.*, vol. 6, no. 1, pp. 488–505, Feb. 2019.

[6] Y. eugin et al., "Building a privacy-preserving smart camera system," in *Proc. Privay Enhancing Technol. Symp.*, 2022, pp. 1–9.

[7] M. Wang, D. Xiao, and Y. Xiang, "Low-cost and confidentiality-preserving multi-image compressed acquisition and separate reconstruction for Internet of Multimedia Things," *IEEE Internet Things J.*, vol. 8, no. 3, pp. 1662–1673, Feb. 2021.

[8] T. Chuman, W. Sirichotedumrong, and H. Kiya, "Encryption-then-compression systems using Grayscale-based image encryption for JPEG images," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 6, pp. 1515–1525, Jun. 2019.

[9] Y. Ding et al., "DeepEDN: A deep-learning-based image encryption and decryption network for Internet of Medical Things," *Internet Things Smart Cities*, vol. 8, no. 3, pp. 1504–1518, 2021.

[10] M. Deruyck, D. Renga, M. Meo, L. Martens, and W. Joseph, "Accounting for the varying supply of solar energy when designing wireless access networks," *IEEE Trans. Green Commun. Netw.*, vol. 2, no. 1, pp. 275–290, Mar. 2018.

[11] G. Mutezo and J. Mulopo, "A review of Africa's transition from fossil fuels to renewable energy using circular economy principles," *Renew. Sustain. Energy Rev.*, vol. 137, Mar. 2021, Art. no. 110609.

[12] C. A. Tracker. "Improvement in Warming Outlook as India and China Move Ahead, But Paris Agreement Gap Still Looms Large." 2017. [Online]. Available: https://climateactiontracker.org/documents/61/CAT_2017-11-15_ImprovementInWarmingOutlook_BriefingPaper.pdf

[13] T. L. Frölicher, E. M. Fischer, and N. Gruber, "Marine heatwaves under global warming," *Nature*, vol. 560, no. 7718, pp. 360–364, 2018.

[14] M. Shah, W. Zhang, H. Hu, and N. Yu, "Paillier cryptosystem based mean value computation for encrypted domain image processing operations," *ACM Trans. Multimedia Comput. Commun. Appl.*, vol. 15, no. 3, pp. 1–21, 2019.

[15] L. Fan, "Image pixelization with differential privacy," in *Proc. Data Appl. Security Privacy*, 2018, pp. 148–162.

[16] D. Münch, A.-K. Grosselfinger, E. Krempel, M. Hebel, and M. Arens, "Data Anonymization for data protection on publicly recorded data," in *Proc. Comput. Vis. Syst.*, 2019, pp. 245–258.

[17] E. von Zezschwitz, S. Ebbinghaus, H. Hussmann, and A. De Luca, "You can't watch this! Privacy-respectful photo browsing on Smartphones," in *Proc. Conf. Human Factory Comput. Syst. Process.*, 2016, pp. 4320–4324.

[18] T. Winkler and B. Rinner, "Security and privacy protection in visual sensor networks: A survey," *ACM Comput. Surveys*, vol. 47, no. 1, pp. 1–42, 2014.

[19] J. Li, N. Wang, L. Zhang, B. Du, and D. Tao, "Recurrent feature reasoning for image Inpainting," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 2020, pp. 7757–7765.

[20] D. Ding, S. Ram, and J. J. Rodríguez, "Image Inpainting using nonlocal texture matching and nonlinear filtering," *IEEE Trans. Image Process.*, vol. 28, no. 4, pp. 1705–1719, Apr. 2019.

[21] R. R. Shetty, M. Fritz, and B. Schiele, "Adversarial scene editing: Automatic object removal from weak supervision," in *Proc. Adv. Neural Inf. Proces. Syst.*, vol. 31, 2018, pp. 7717–7727.

[22] J. Cao, B. Liu, Y. Wen, R. Xie, and L. Song, "Personalized and invertible face de-identification by disentangled identity information manipulation," in *Proc. IEEE Int. Conf. Comput. Vis.*, Oct. 2021, pp. 3334–3342.

[23] H.-P. Wang, T. Orekondy, and M. Fritz, "InfoScrub: Towards attribute privacy by targeted obfuscation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR) Workshops*, Jun. 2021, pp. 3281–3289.

[24] E. T. Hassan, R. Hasan, P. Shaffer, D. Crandall, and A. Kapadia, "Cartooning for enhanced privacy in lifelogging and streaming videos," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Workshops*, Jul. 2017, pp. 1333–1342.
[25] M. A. Taha, N. Sidaty, W. Hamidouche, O. Dforges, J. Vanne, and M. Viitanen, "End-to-end real-time ROI-based encryption in HEVC videos," in *Proc. Eur. Signal Proces. Conf.*, 2018, pp. 171–175.
[26] J. He et al., "PUPPIES: Transformation-supported personalized privacy preserving partial image sharing," in *Proc. Annu. IEEE/IFIP Int. Conf. Depend. Syst. Netw.*, 2016, pp. 359–370.
[27] M. Saini, P. K. Atrey, S. Mehrotra, and M. Kankanhalli, "W3-Privacy: Understanding what, when, and where inference channels in multi-camera surveillance video," *Multimedia Tools Appl.*, vol. 68, no. 1, pp. 135–158, 2014.
[28] J.-Y. Zhu, T. Park, P. Isola, and A. A. Efros, "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2017, pp. 2242–2251.
[29] H. Wu et al., "PECAM: Privacy-enhanced video streaming and Analytics via securely-reversible transformation," in *Proc. Annu. Int. Conf. Mobile Comput. Netw.*, 2021, pp. 229–241.
[30] X. Chai, Y. Wang, X. Chen, Z. Gan, and Y. Zhang, "TPE-GAN: Thumbnail preserving encryption based on GAN with key," *IEEE Signal Process. Lett.*, vol. 29, pp. 972–976, 2022.
[31] S. Çiftçi, A. O. Akyüz, and T. Ebrahimi, "A reliable and reversible image privacy protection based on false Colors," *IEEE Trans. Multimedia*, vol. 20, no. 1, pp. 68–81, Jan. 2018.
[32] K. Tajik et al., "Balancing image privacy and usability with thumbnail-preserving encryption," in *Proc. Symp. Netw. Distrib. Syst. Security*, 2019, pp. 1–9.
[33] R. Zhao, Y. Zhang, X. Xiao, X. Ye, and R. Lan, "TPE2: Three-pixel exact thumbnail-preserving image encryption," *Signal Process.*, vol. 183, Jun. 2021, Art. no. 108019.
[34] Y. Zhang, R. Zhao, X. Xiao, R. Lan, Z. Liu, and X. Zhang, "HF-TPE: High-fidelity thumbnail-preserving encryption," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 32, no. 3, pp. 947–961, Mar. 2022.
[35] Y. Zhang, R. Zhao, Y. Zhang, R. Lan, and X. Chai, "High-efficiency and visual-usability image encryption based on thumbnail preserving and chaotic system," *J. King Saud Univ. Comput. Inf. Sci.*, vol. 34, no. 6, pp. 2993–3010, 2022.
[36] C. V. Wright, W.-C. Feng, and F. Liu, "Thumbnail-preserving encryption for JPEG," in *Proc. ACM Workshop Inf. Hiding Multimedia Security*, 2015, pp. 141–146.
[37] W. Li, Y. Yan, and N. Yu, "Breaking row-column shuffle based image cipher," in *Proc. ACM Int. Conf. Multimedia*, 2012, pp. 1097–1100.
[38] A. Jolfaei, X.-W. Wu, and V. Muthukkumarasamy, "On the security of permutation-only image encryption schemes," *IEEE Trans. Inf. Forensic Security*, vol. 11, no. 2, pp. 235–246, Feb. 2016.
[39] M. Bellare, T. Ristenpart, P. Rogaway, and T. Stegers, *Format-Preserving Encryption* (Lecture Notes in Computer Science). Cham, Switzerland: Springer, Aug. 2009, pp. 295–312.
[40] Z. Hua, Z. Zhu, Y. Chen, and Y. Li, "Color image encryption using orthogonal latin squares and a new 2D chaotic system," *Nonlinear Dyn.*, vol. 104, no. 4, pp. 1–18, 2021.
[41] G. Elkoumy et al., "Privacy and confidentiality in process mining: Threats and research challenges," *ACM Trans. Manag. Inf. Syst.*, vol. 13, no. 1, pp. 1–17, 2021.
[42] M. Langheinrich, "Privacy by design—Principles of privacy-aware ubiquitous systems," in *Proc. Int. Conf. Ubiquitous Comput. Commun.*, 2001, pp. 273–291.
[43] R. Zhao, Y. Zhang, Y. Nan, W. Wen, X. Chai, and R. Lan, "Primitively visually meaningful image encryption: A new paradigm," *Inf. Sci.*, vol. 613, pp. 628–648, Oct. 2022.
[44] Y. Zhang, W. Zhou, R. Zhao, X. Zhang, and X. Cao, "F-TPE: Flexible thumbnail-preserving encryption based on multi-pixel sum-preserving encryption," *IEEE Trans. Multimedia*, early access, Aug. 19, 2022, doi: 10.1109/TMM.2022.3200310.
[45] V. Le, J. Brandt, Z. Lin, L. Bourdev, and T. S. Huang, "Interactive facial feature Localization," in *Proc. Comput. Vis. ECCV*, 2012, pp. 679–692.
[46] D. R. I. M. Setiadi, "PSNR vs SSIM: Imperceptibility quality assessment for image steganography," *Multimedia Tools Appl.*, vol. 80, no. 6, pp. 8423–8444, 2021.
[47] K. Muhammad, R. Hamza, J. Ahmad, J. Lloret, H. Wang, and S. W. Baik, "Secure surveillance framework for IoT systems using probabilistic image encryption," *IEEE Trans. Ind. Informat.*, vol. 14, no. 8, pp. 3679–3689, Aug. 2018.

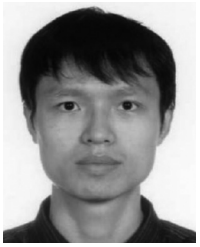
Ruoyu Zhao received the B.S. degree in computer science and technology from the School of Software, Zhengzhou University, Zhengzhou, China, in June 2019, and the M.S. degree in cyberspace security from the College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing, China, in April 2022, where he is currently pursuing the Ph.D. degree in cyberspace security.

His research interests include multimedia security and privacy.

Yushu Zhang (Member, IEEE) received the Ph.D. degree from the College of Computer Science, Chongqing University, Chongqing, China, in December 2014.

He held various research positions with the City University of Hong Kong, Hong Kong; Southwest University, Chongqing; the University of Macau, Macau, China; and Deakin University, Geelong, VIC, Australia. He is currently a Professor with the College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing, China. His research interests include multimedia security, artificial intelligence security, and blockchain. He has published over 150 refereed journal articles and conference papers in these areas.

Prof. Zhang is an Editor of *Information Sciences*, *Journal of King Saud University—Computer and Information Sciences*, and *Signal Processing*.



Rushi Lan received the B.S. degree in information and computing science and the M.S. degree in applied mathematics from Nanjing University of Information Science and Technology, Nanjing, China, in 2008 and 2011, respectively, and the Ph.D. degree in software engineering from the University of Macau, Macau, China, in 2016.

He is currently an Associate Professor with the School of Computer Science and Information Security, Guilin University of Electronic Technology, Guilin, China. His research interests include image classification, image denoising, and metric learning.

Zhongyun Hua (Member, IEEE) received the B.S. degree in software engineering from Chongqing University, Chongqing, China, in 2011, and the M.S. and Ph.D. degrees in software engineering from the University of Macau, Macau, China, in 2013 and 2016, respectively.

He is currently an Associate Professor with the School of Computer Science and Technology, Harbin Institute of Technology (Shenzhen), Shenzhen, China. His research interests include chaotic system, chaos-based applications, data hiding, and multimedia security. He has published more than 60 papers on the subject, receiving more than 3800 citations.

Yong Xiang (Senior Member, IEEE) received the Ph.D. degree in electrical and electronic engineering from the University of Melbourne, Melbourne, VIC, Australia, in 2003.

He is a Professor with the School of Information Technology, Deakin University, Burwood, VIC, Australia. His research interests include distributed computing, cybersecurity and privacy, machine learning and AI, and communications technologies. He has published seven monographs, over 210 refereed journal articles, and over 100 conference papers in these areas.

Prof. Xiang is a Senior Area Editor of IEEE SIGNAL PROCESSING LETTERS and an Associate Editor of IEEE COMMUNICATIONS SURVEYS AND TUTORIALS and *Computer Standards and Interfaces*. He was an Associate Editor of IEEE SIGNAL PROCESSING LETTERS and IEEE ACCESS, and a Guest Editor of IEEE TRANSACTIONS ON INDUSTRIAL INFORMATICS and IEEE MULTIMEDIA. He has served as the honorary chair, general chair, program chair, TPC chair, symposium chair, and track chair for many conferences, and was invited to give keynotes at numerous international conferences.