Signal Processing 183 (2021) 108019

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

TPE2: Three-Pixel Exact Thumbnail-Preserving Image Encryption

Ruoyu Zhao[a], Yushu Zhang[a,b,*], Xiangli Xiao[a], Xi Ye[a], Rushi Lan[c]

[a] College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China
[b] Guangxi Key Laboratory of Trusted Software, Guilin University of Electronic Technology, Guilin 541004, China
[c] Guangxi Key Laboratory of Image and Graphic Intelligent Processing, Guilin University of Electronic Technology, Guilin 541004, China

ARTICLE INFO

Article history:
Received 31 October 2020
Revised 30 December 2020
Accepted 27 January 2021
Available online 30 January 2021

2020 MSC:
00-01
99-00

Keywords:
Image usability
Cloud storage
Privacy protection

ABSTRACT

Facing a large number of personal images and local devices with limited resources, the cloud service plays an increasingly important role in image storage. However, there is no doubt that this will raise a lot of privacy concerns if the image is not under the direct control of the owner. Simply encrypting images utilizing traditional encryption schemes protects the privacy of them, but it sacrifices the usability of image content. Resolving the tension between the usability and privacy risks of images is a critical issue for users. Recently, Tajik et al. proposed the exact thumbnail preserving image encryption (TPE) scheme as a group of two pixels to balance image privacy and usability in the cloud. However, the scheme utilizes the Markov chain to prove the security. The fewer pixels in the group, the worse the connectivity of the chain state. Motivated by this, we further extend the TPE scheme to propose a three-pixel exact TPE scheme, called TPE2. The scheme achieves exact TPE as a group of three pixels. The experimental evaluations show that the image encrypted by the scheme has a good image quality and a satisfactory balance usability and privacy.

1. Introduction

Over the past decade, taking images is no longer the privilege of some people with professional cameras. Anyone can utilize their mobile phones to take images anytime and anywhere with the rapid popularity of smartphones with high-definition cameras. It is reported that 1.42 trillion images were taken in 2019 and the annual growth rate is considerable fast that it will take more than 1.56 trillion images in 2022 [1]. The images often inadvertently record people's daily lives, which may lead to leakage of too much sensitive information [2–4] such as religion, ethnicity, and social class. An online survey [5] shows that the majority of respondents (88.6%) believed that some private images are stored on their phones. Moreover, although the definition of private images is subjective, almost all respondents consider images containing family members or personal identification to be privacy. A considerable number of images taken contain both of them. As a result, people often consciously protect their image privacy, which is very easy on local devices, simply by preventing others from browsing their personal images.

However, more and more people are migrating images from local devices to the cloud (e.g., iCloud, OneDrive, and Dropbox) due to the limited resources of local devices (e.g., storage space) and the inconvenience of using images across devices. Even mainstream cloud services provide the option of automatically synchronizing local images to the cloud, which makes it easier for users to enjoy cloud services. The popularity of cloud services allows users to browse and download images from any internet-connected device, anytime, anywhere, and easily display their images online (especially on social platforms). For example, according to statistics [6], 1022 images are uploaded to Instagram every second. Placing images in the cloud undoubtedly greatly reduces the storage burden on local devices. However, these images are no longer directly under the physical control of the image owner and therefore it greatly increases the difficult for the owner to protect the privacy of images.

Enck et al. [7] surveyed 30 popular apps and found that two-thirds of them abuse the users' privacy data. Moreover, according to the survey conducted by Li et al. [5], most people (87.5%) are not sure whether the application they have installed can access the users' album without the users' knowledge. By default, Facebook collects vast amounts of user privacy data. Although Facebook data applying policy states[1] that the data will not be sold to others

* Corresponding author at: College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China.
E-mail address: yushu@nuaa.edu.cn (Y. Zhang).

[1] https://www.facebook.com/about/privacy/update

https://doi.org/10.1016/j.sigpro.2021.108019
0165-1684/

Fig. 1. Top: The images uploaded to the cloud are plaintext images, which can not only be browsed by users, but also obtained by illegal third parties. **Middle:** The images uploaded to the cloud are encrypted by traditional encryption schemes. Although it prevents illegal third parties from learning the privacy of images, it also prevents users from learning information. **Below:** The images uploaded to the cloud are encrypted by the TPE scheme. It prevents illegal third parties from learning image privacy, while users can also learn information by previewing encrypted images.

without the users' consent, apparently it has not kept its promise. This is particularly evident in the case of Cambridge Analytica, which seized huge amounts of user data from Facebook without users' permission [8]. Meanwhile, Facebook was found accidentally to tag images uploaded to the platform without the user's knowledge in 2019 [9]. In addition, the 2014 iCloud image leak was one of the worst cloud privacy breaches to data, with hundreds of Hollywood celebrities' private images leaked by hackers [10].

Intuitively, image privacy problem can be simply solved by traditional image encryption technologies (e.g., [11–13]). The users' images are encrypted before stored in the cloud to protect privacy. However, traditional encryption technologies make images lose their usability and visibility in cloud scenarios. For example, users are no longer able to preview and select images in the cloud, nor can they organize and manage images in the cloud. All operations can only be performed after downloading and decrypting ciphertext images. In other words, traditional encryption deprives images of content-based usability. Although privacy and usability of images are valuable to users, traditional encryption cannot achieve a satisfactory balance between the two. Even for many ordinary users, though they know the benefits of protecting image privacy, they still give up encrypted images for usability in the cloud. This also explains why it is more common in the real world to store plaintext images in the cloud than ciphertext ones. Therefore, in order to please users, the privacy protection of images should not be at the expense of the usability of images in the cloud.

To achieve a satisfactory balance between image privacy and usability in the cloud, Wright et al. [14] proposed the concept of thumbnail-preserving encryption (TPE), i.e., ciphertext images encrypted present the same visual content as low-resolution versions of the original ones. Specifically, TPE preserves rough information that is large than the thumbnail block and erases precise information that is smaller than the one. The image owner is able to identify images based on the rough visual information preserved by ciphertext images. The image owner can obtain usability, while others do not possess this ability since they have not browsed original images in advance. Therefore, the TPE scheme can achieve a satisfactory balance image privacy protection and usability, as shown in Fig. 1.

Wright et al. [14] designed the first TPE scheme, which applies permutation-only encryption and thus ciphertext images can exactly preserve the original image thumbnail. However, ciphertext images produced by this scheme leak much more information than the thumbnails. In addition, long-term studies have shown that the security of permutation-only image encryption is insufficient in some aspects [15–17]. For example, jolfaei et al. proved [15] that the plaintext can be completely recovered by chosen plaintext attacks regardless of the encryption structure. Subsequently, Marohn et al. [18] proposed two approximate TPE schemes. Although these two schemes are claimed to be secure by the authors, ciphertext images reveal more information than the thumbnails themselves and the quality of the decrypted images is poor. Recently, Tajik et al. [19] designed the first TPE scheme to achieve nonce-respecting (NR) security by exploiting format-preserving encryption [20]. In this scheme, Markov chain is applied for security analysis and the ciphertext images can preserve exact thumbnails. However, this TPE scheme can only substitute encryption with two pixels as a group. As mentioned by Tajik et al. [19] in the future directions, a more effective substitution functions can be considered in the the subsequent extensibility work, i.e., substitution encryption for more pixels, which will increase the connectivity of the Markov chain.

With this as the motivation, this work conducts in-depth research on the previous TPE schemes and proposes a more effective substitution encryption function to increase the connectivity of the Markov chain, i.e., the TPE2 scheme. This scheme encrypts the image with three pixels as a group and the results obtained can exactly preserve the original image thumbnail. In addition, this scheme achieves NR security.

Contribution: The main contributions of this work can be highlighted as follows:

- We deeply analyze the relationship between the sum of three pixels and the number of pixel groups and propose the formula to directly calculate the number of pixel groups based on the sum of three pixels. Meanwhile, the formula is combined with the sum-preserving encryption algorithm and the format-preserving encryption algorithm to propose the TPE2 scheme, which is the first scheme to extend the exact TPE to a group of three pixels for substitution encryption.
- The proposed TPE2 scheme is compatible with existing cloud storage services and the encrypted images can be uploaded to the existing cloud in a way of ensuring privacy and usability without any modifications to the cloud system.
- The analysis indicates that the TPE2 scheme achieves NR security. Meanwhile, the parameters of face detection algorithm failure and the usability of ciphertext images under this parameters are evaluated through experiments. The evaluation results show that the scheme can achieve a good balance between image usability and privacy.

The rest of this paper is organized in the following manner. Section 2 reviews some potential solutions that might ease the conflict between image privacy and usability and points out their shortcomings. Section 3 briefly introduces some cryptographic primitives utilized in the TPE2 scheme. Section 4 gives the threat model, goals, and assumptions. Section 5 presents the concrete construction and security analysis of the TPE2 scheme. The empirical evaluations are given in Section 6. Section 7 concludes this work and discusses prospects.

2. Potential solutions

A potential solution to ease the conflict between privacy and usability of images in the cloud is image annotation [21]. This scheme exploits computer vision or manual method to annotate

the visual content of images before encryption (i.e., several keywords are assigned to the image) and then annotations are encrypted and uploaded to the cloud together with ciphertext images. The scheme ensures the privacy of images in the cloud since illegal third parties cannot learn information from ciphertext images. Meanwhile, users can identify images through annotation information and then select, organize, and manage images, i.e., usability is achieved. However, there are serious problems with this solution. First, how to annotate the image? Although many excellent image annotation technologies (e.g., [22–24]) have been proposed due to the development of machine learning technology and even there are some technologies (e.g., [25,26]) for resource-limited devices such as phone, the information annotated by these technologies may not be helpful to users. Meanwhile, it is impractical to impose the task of annotating images on users. Second, similar or even the same annotation will inevitably appear when there are emerging many images, which makes it impossible for users to distinguish images by annotations. Third, this scheme can not be applied in mainstream cloud services, such as OneDrive, iCloud, and Google Drive, since they do not support image annotation uploading.

Image retrieval or searchable encryption is also considered as a potential solution to mitigate the conflict between privacy concerns and usability of images (e.g., [27–29]). In this solution, the features (e.g., SIFT [30] and SURF [31]) or keywords of plaintext are extracted as indexes by algorithms before image encryption and then the indexes and images are encrypted and uploaded to the cloud. Users obtain desired images based on the guarantee of image privacy by exacting or similarity searching on the index. However, the problem with this solution is how to prevent the untrusted cloud from learning the relationship between retrieved information and images. Although this problem can be solved by homomorphic encryption to a certain extent, its computational cost is too expensive to be applied in practice. In addition, users may not know what images they want until they browse the visual content of the image and therefore they cannot utilize the retrieval service to get images.

Some visual psychology studies have opened up ideas for solutions to balance image privacy and usability in the cloud. Gregory's [32] research shows that people have the ability to recognize distorted images (especially faces) they have browsed. Later, Snodgrass's [33] research further shows that this ability would be stronger if the images were taken by themselves. Research by Guillaume et al. [34] indicates that people can remember the key features of images only after a short glance. These remembered features can be called prior knowledge. The experiment conducted by Denning et al. [35] demonstrates that people still have the ability to recognize images based on broken visual information combined with prior knowledge, after a long time of browsing images. These studies show that image degradation can be regarded as a one-way function. The two ends of the function are the original image and the degraded version of the original one (i.e., the low-resolution image). It is easy for legitimate users to visually restore the image by browsing the degraded version of the image in conjunction with prior knowledge, but it is difficult for illegal third parties.

Some solutions have been proposed to protect image privacy and preserve image usability by exploiting image degradation such as pixelate, crystallize, and obfuscation (e.g., [36–38]). On the one hand, images processed by such solutions contain enough visual information to enable the image owner to extract the features of images and thus preserving usability. On the other hand, processed images make it impossible for illegal third parties to learn image privacy. However, these solutions are not suitable for cloud storage services, since they are mainly aimed at privacy concerns caused by image sharing and do not consider restoring original images. Recently, Sun et al. [39] proposed a reversible cryptographic mosaic technique by permutations, which utilized cryptography to render the visual effect of ciphertext images to be the low-resolution version of plaintext ones. However, this solution only utilizes the permutation-only encryption, which as mentioned above, the security of this encryption is insufficient.

Region of interest (ROI) encryption or partial encryption is also often used as a solution to balance image privacy and usability (e.g., [40,41]). Such solutions select the privacy region in the image by algorithm or by hand, which is called the secret part, and then encrypt it. The rest part is not processed, which is called the public part. Theoretically, the illegal third party cannot learn the privacy in the image and the public part of the image can help the image owner to understand and recognize the image. However, the main problem of the solution is how to select the secret part of the image. The algorithm may not be able to correctly divide the secret part of the image due to the subjectivity of privacy. Meanwhile, it is unrealistic for the image owner to manually partition.

Table 1
Notations in cryptosystems.

Notation	Description
$\mathbb{Z}_{d+1}$	Non-negative numbers not greater than d
$\mathcal{M}$	Message space
M	Plaintext
C	Ciphertext
T	Nonce, $T \in \{0,1\}^*$
K	Symmetric key, $K \in \{0,1\}^\lambda$
Enc_K	Encryption algorithm
Dec_K	Decryption algorithm
Φ	Function that preserves specific format on $\mathcal{M}$
F	Function : $\{0,1\}^* \times \mathcal{M} \to \mathcal{M}$
$\mathcal{F}_\Phi$	The set of functions F
SPEnc_K	Sum-preserving encryption algorithm
SPDec_K	Sum-preserving decryption algorithm
Φ_{sum}	Function that preserves the sum on $\mathcal{M}$

3. Preliminaries

In this section, we mainly review some cryptographic primitives related to this paper. For brevity, the notations utilized in the section are shown in Table 1.

3.1. Thumbnail-preserving encryption

TPE means that the ciphertext image is a low-resolution version (i.e., thumbnail) of the plaintext one. Specifically, ciphertext images and plaintext ones have the same dimensions and thumbnails. On the one hand, as described above, the image owner is able to accurately recognize and preview images based on these low-resolution ciphertext images, i.e., preserving images' usability. On the other hand, although ciphertext images present some image information in the visual effect, the leaked information does not exceed plaintext thumbnails.

The generation of thumbnails is based on blocks. First, the image is divided into $b \times b$ blocks and then the average value of the sum of pixels in each block is calculated, which is the pixel values in the thumbnails corresponding to the block. This means that the sums of the pixels of the blocks in the image encrypted by the exact TPE scheme and the corresponding blocks in the original image is the same.

3.2. Nonce-based encryption

The TPE scheme is a specific type of probabilistic encryption. The probabilistic encryption means that encrypting the same plaintext multiple times will produce different ciphertexts. The nonce should be selected in the process of probabilistic encryption and

then encrypted. It is known that the ciphertext of probabilistic encryption will generate ciphertext expansion in order to resist chosen plaintext attacks. However, the image encrypted by the TPE scheme should have the same dimension as the original one. Meanwhile, the TPE scheme cannot requires cloud storage services to store any encryption-related auxiliary data other than images themselves, since existing cloud services do not support this operation. In other words, the TPE scheme should prevent ciphertext expansion. In fact, the image itself has unique identifiers (e.g., filename and shooting time) that can act as a unique nonce during encryption.

The (nonce-based) encryption algorithm takes K, T, and M as input, rutruns C as output; The (nonce-based) decryption algorithm takes K, T, and C as input, returns M as output, which as shown in following equations.

$$\mathrm{Enc}_K(T, M) = C$$
$$\mathrm{Dec}_K(T, \mathrm{Enc}_K(T, M)) = M$$

3.3. Format-preserving encryption

The TPE scheme is also a specific type of format preserving encryption (FPE). FPE refers to the encryption of the plaintext with some specified formats into the ciphertext with the same formats [20]. For the TPE scheme, it is equivalent to choosing a set of images, that is $\mathcal{M}$, $M \in \mathcal{M}$, and $\Phi(M)$ has the same dimensions and low-resolution thumbnails as M. An encryption scheme is considered as Φ-preserving if it satisfies the following equation:

$$\mathrm{Enc}_K(T, M) \in \mathcal{M}$$
$$\Phi(\mathrm{Enc}_K(T, M)) = \Phi(M)$$

If $\{C \mid \Phi(C) = \Phi(M)\}$ is treated as the slices of $\mathcal{M}$, i.e., each item of the $\mathcal{M}$ belongs to the slice of the format space, then the Φ-preserving encryption definition is exactly the same as the original FPE definition [19].

Bellare et al. [20] given some definitions regarding FPE security, as shown in the following:

Definition 1. The function F is considered as Φ-preserving, that is $\Phi(F(T, M)) = \Phi(M)$. The FPE scheme is Pseudo Random Permutation (PRP) security (it is analogous to chosen plaintext attack (CPA)-security), if it meets the following property for all Probabilistic Polynomial Time (PPT) oracle machines $\mathcal{A}$ [20]:

$$\left| \Pr_{K \leftarrow \{0,1\}^{\lambda}} \left[\mathcal{A}^{\mathrm{Enc}_K(\cdot,\cdot)}(\lambda) = 1 \right] - \Pr_{\mathrm{F} \leftarrow \mathcal{F}_{\Phi}} \left[\mathcal{A}^{\mathrm{F}(\cdot,\cdot)}(\lambda) = 1 \right] \right|$$

is negligible in λ.

Definition 2. $\mathcal{A}$ is Nonce-Respecting (NR) if it never makes two oracle calls with the same first argument (i.e., nonce T) [19].

The FPE scheme is NR security if the scheme satisfies Definition 1 but only under the condition of Definition 2. As mentioned above, each image has a natural and unique identifier that can be utilized to act as T. Under the premise that no two images have the same T, the NR security has the same guarantee as PRP security. In other words, ciphertext images are indistinguishable from randomly chosen images with the same thumbnail, as shown in Fig. 2.

3.4. Sum-preserving encryption

It can be known from the above that the core of the TPE scheme is to keep the sum of pixel values in each thumbnail block unchanged before and after encryption. Therefore, the TPE is a kind of sum-preserving encryption (SPE). The properties of the SPE are as follows:

$$\mathcal{M} = (\mathbb{Z}_{d+1})^n$$
$$\vec{v} = (v_1, \ldots, v_n) \in \mathcal{M}$$
$$\Phi_{sum}(\vec{v}) = \sum_{i=1}^{n} v_i$$
$$\mathrm{SPEnc}_K(T, \vec{v}) = \vec{c}$$
$$\vec{c} = \{c_1, \ldots, c_n\} \in \mathcal{M}$$
$$\mathrm{SPDec}_K(T, \vec{c}) = \vec{v}$$
$$\Phi_{sum}(\mathrm{SPEnc}_K(T, \vec{v})) = \Phi_{sum}(\vec{v})$$

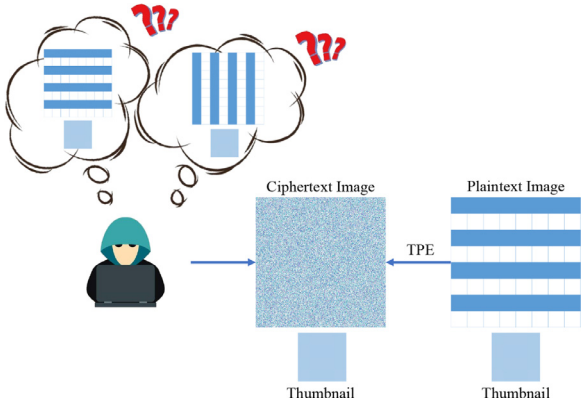

Fig. 2. The ciphertext image is indistinguishable from randomly chosen images that share the same thumbnails.

4. Threat model, goals, and assumptions

In this paper, we focus on two types of attacks, i.e., insider attacks and outsider attacks.

Insider attacks: There are three main security issues within cloud storage services. The first security issue is that cloud service providers themselves may learn the information in user images to mine the potential value of users (e.g., accurate recommendation of advertisement and user portrait). The second security issue is that it is difficult for users to completely delete images with privacy issues after they are uploaded to the cloud. Even if users delete the current service's images, other copies are still on cloud services, because cloud storage services usually store multiple data backups on multiple servers. The third security issue is that cloud storage employees may be rogue. They may sell users' privacy to third parties for economic benefits.

Outsider attacks: The main outsider threat to cloud services is the illegal third party (e.g., hacker), which is an active adversary. It can intercept between cloud services and client, attack cloud servers, and even associate with the cloud.

The goal of this paper is to design an encryption scheme that can protect the privacy of users' images (as opposed to the two types of attacks mentioned above). At the same time, without changing existing cloud services, users can obtain usability (e.g., previewing) only through the information of images themselves (i.e., visual content).

The above goal implies several limitations of the scheme, which makes it very challenging. First, the scheme needs to be compatible with existing cloud services, i.e., the scheme needs to be deployed without any challenges to cloud services (e.g., usage, software, and backend). Second, images are encrypted during transmission and storage to the cloud while only users can decrypt and obtain plaintext images. In other words, the usability comes from ciphertext images themselves, not from other additional information. Third, users need to be able to control the privacy granularity of ciphertext images exposure due to the subjectivity of privacy.

We make the following assumptions about the trust of all parties involved in the cloud service scenario. *a*) The local hardware

Table 2
Notations in TPE2 scheme construction.

Notation	Description
$\mathcal{M}$	Format space, $\mathcal{M} \in (\mathbb{Z}_{d+1})^n$
M	Plaintext image
C	Ciphertext image
s	Sum of elements in the vector
c	Number of channels in the image
b	The dimension of the thumbnail block
w, h	The width and height of the image, respectively
w', h'	w/b and h/b, respectively
k	The kth channel of the image, $k \in \{1, \ldots, c\}$
j	The column number of the block, $j \in \{1, \ldots, w'\}$
i	The row number of the block, $i \in \{1, \ldots, h'\}$
$I_b[k, i, j]$	The (i, j)th block of the kth channel in image I
b2v$(I_b[k, i, j])$	Converting two-dimensional block into one-dimensional vector
$\Phi^{-1}_{sum}(s)$	The set of vectors with the same sum s in $\mathcal{M}$
$\lvert\Phi^{-1}_{sum}(s)\rvert$	The number of vectors in $\Phi^{-1}_{sum}(s)$
$\lvert\lvert$	Connector
sub	Part of nonce in substitution encryption
per	Part of nonce in permutation encryption
SN	The sequence number of the vector
rank$_s(.)$	Function that find the SN of the vector in $\Phi^{-1}_{sum}(s)$
rank$_s^{-1}(.)$	Function that find the vector corresponding to the SN in $\Phi^{-1}_{sum}(s)$
params(I)	$(c, w, h) = params(I)$, which finds the parameters of the image I

and software on the user's client, such as smartphones and personal computers, including systems and applications, are trusted. *b*) The cloud services are semi-honest parties (i.e., insider attacks). They can store and download images correctly, but they are curious about the privacy information, original images, and the key. *c*) The illegal third parties are simply not trusted (that is, outsider attacks).

5. TPE2 construction

5.1. Overview

In this scheme, we consider the following types of images:

- Each image consists of one or more *channels*, such as grayscale, RGB, CMKY.
- Each image has a unique set of public metadata that can be thought of as *nonce* (i.e., T).
- Each image channel is a matrix composed of pixels, whose values/intensities are taken from $\{0, \ldots, d\}$ (often, $d = 255$, and in the following, d is 255 unless otherwise specified).

For the sake of convenience, the notations utilized in the TPE2 scheme and their corresponding descriptions are shown in Table 2.

5.2. Description of scheme

In the Section 3, it is stated that the core of the TPE2 scheme lies in keeping the sum of pixel values in each thumbnail block unchanged before and after encryption. Since the image is composed of multiple channels, the sum of pixel values in the ciphertext image block can be guaranteed to keep unchanged by ensuring the sum of the corresponding block values in each channel. Therefore, this scheme splits the image into multiple channels and then independently encrypts each channel by TPE2.

The SPE is a natural way to construct the TPE2 scheme. Specifically, the SPE algorithm is separately applied for each thumbnail block in channel to preserve the thumbnail of the original image in the ciphertext image. The general TPE2 encryption algorithm is shown in Algorithm 1. The general TPE2 decryption algorithm is shown in Algorithm 2. Tajik et al. [19] have proved that as long as the encryption and decryption algorithms of the SPE are NR-secure scheme, the TPE scheme constructed by the SPE algorithms is a NR-secure TPE.

Algorithm 1 The general TPE2 encryption algorithm.

Input: T, b, K, and M
Output: C

$(c, w, h) = \text{params}(M)$
for $k = 1 : c$ **do**
  for $i = 1 : h'$ **do**
    for $j = 1 : w'$ **do**
      $T' = T||k||i||j$
      $\vec{v}_{mb}$= b2v$(M_b[k, i, j])$
      $C_b[k, i, j] = \text{SPEnc}_K(T', \vec{v}_{mb})$
    end for
  end for
end for
return C

Algorithm 2 The general TPE2 decryption algorithm.

Input: T, b, K, and C
Output: M

$(c, w, h) = \text{params}(C)$
for $k = 1 : c$ **do**
  for $i = 1 : h'$ **do**
    for $j = 1 : w'$ **do**
      $T' = T||k||i||j$
      $\vec{v}_{mb}$= b2v$(C_b[k, i, j])$
      $M_b[k, i, j] = \text{SPDec}_K(T', \vec{v}_{mb})$
    end for
  end for
end for
return M

However, construing a SPE algorithm is a very challenging task because it requires that each element value in the ciphertext vector be a member of $\mathbb{Z}_{d+1}$ and the sum of the elements of the vector be preserved. The rank-then-encipher (RTE) proposed by Bellare et al. [20] is a good solution. The RTE scheme first enumerates all elements (e.g., $\Phi^{-1}_{sum}(s)$) that have the same format on $\mathcal{M}$ and assigns them a unique SN. Then, the plaintext SN can be obtained by utilizing function rank$_s(.)$. After encrypting the SN (within the scope of $\mathbb{Z}_{|\Phi^{-1}_{sum}(s)|}$), the encrypted SN is converted into ciphertext

by utilizing function $\text{rank}_s^{-1}(.)$. The SPE algorithm constructed by RTE scheme is shown in Algorithm 3. The step 10 of this algorithm

Algorithm 3 RTE-SPE encryption algorithm.

Input: T', K, and $\vec{v} = (v_1, \ldots, v_n) \in \mathcal{M}$
Output: $\vec{vc} = (y_1, \ldots, y_n) \in \mathcal{M}$ /* $\sum_{i=1}^{n} y_i = \sum_{i=1}^{n} v_i$ */
1: $s = \sum_{i=1}^{n} v_i$
2: $\mathcal{N} = \emptyset$
3: $\Phi_{sum}^{-1}(s) = \left\{\vec{x} \,|\, \vec{x} = (x_1, \ldots, x_n) \in \mathcal{M}, s = \sum_{i=1}^{n} x_i\right\}$
4: **for** $num = 0 : |\Phi_{sum}^{-1}(s)| - 1$ **do**
5: $\quad \vec{z} \in (\Phi_{sum}^{-1}(s) - \mathcal{N})$
6: $\quad$ Set the SN of $\vec{z}$ to be num
7: $\quad \mathcal{N} = \mathcal{N} \cup \vec{z}$
8: **end for**
9: $r = \text{rank}_s(\vec{v})$ /* the SN of $\vec{v}$ */
10: $cr = \text{encrypt } r \text{ by } \text{Enc}_K(T', r)$, $cr \in \mathbb{Z}_{|\Phi_{sum}^{-1}(s)|}$
11: $\vec{vc} = \text{rank}_s^{-1}(cr)$ /* the cr corresponds to vector in set $\Phi_{sum}^{-1}(s)$ */
12: **return** $\vec{vc}$

is changed to cr **= decrypt** r **by** $\text{Dec}_K(T', r)$, which is the corresponding decryption algorithm.

The pixels in the block are divided into groups as the computational complexity of directly encrypting the pixels in the thumbnail block is very high. In the TPE2 scheme, the three pixels are divided into a group, which can be regarded as one-dimensional vector. Then the pixel group is encrypted by RTE-SPE algorithm. It is necessary to know the number of pixel groups with the same sum when encrypting, i.e., $|\Phi_{sum}^{-1}(s)|$, so as to assign SN to the pixel group and prevent the encrypted SN from crossing the boundary.

Lemma 1. *The function* $|\Phi_{sum}^{-1}(s)|$ *is shown below when three pixels are divided into a group.*

$$|\Phi_{sum}^{-1}(s)| = \begin{cases} \text{sum}(0, 1, l_1) & 0 \le s < l_2 \\ |\Phi_{sum}^{-1}(d)| + sum(l_2, -2, l_3) & l_2 \le s \le l_4 \\ \text{sum}(0, 1, l_5) & \text{otherwise} \end{cases}$$

where $l_1 = s + 1$. $l_2 = d + 1$, $l_3 = s - d$, $l_4 = 2 \times l_2 - 1$, $l_5 = 3 \times d - s + 1$, *function* $\text{sum}(i, j, k)$ *is shown in Algorithm 4.*

Algorithm 4 sum(x,y,z).

Input: x, y, and z
Output: num_s
1: $num_s = 0$
2: $q = \text{x}$
3: **for** $m = 1 : z$ **do**
4: $\quad q = q + y$
5: $\quad num_s = num_s + q$
6: **end for**
7: **return** num_s

Proof 1. The function $|\Phi_{sum}^{-1}(s)|$ reaches its maximum value, when the sum s of the pixel group is $\frac{3}{2}d$. Meanwhile, the symmetry axis of the function image is at s= $\frac{3}{2}d$, i.e., $|\Phi_{sum}^{-1}(s_1)| = |\Phi_{sum}^{-1}(3d - s_1)|$ $(s_1 \le \frac{3}{2}d)$. Therefore, the lemma only needs to prove the case of $0 \le s \le \frac{3}{2}d$. For simplicity, the three pixels of the group are called a, b, and c, respectively. When there are only two pixels (b, c), it is easy to get that when $b + c \le d$, (b, c) has $b + c + 1$ cases, otherwise, (b, c) has $2d - b - c + 1$ cases. When $0 \le s \le d$, a can be any value in $\mathbb{Z}_s$, $b + c = s - a$, and $0 \le s - a \le d$. When a is determined, there are $s - a + 1$ cases in (b, c), thus $|\Phi_{sum}^{-1}(s)| = \sum_{a=0}^{s}(s - a + 1) = \text{sum}(0, 1, l_1)$. Similarly, when $d \le s \le \frac{3}{2}d$ and $0 \le a < s - d$, there are $|\Phi_{sum}^{-1}(d)| - \text{sum}(0, 1, s - d)$ cases. When $d \le s \le \frac{3}{2}d$ and $s - d \le a \le d$, there are $\text{sum}(d + 1, -1, s - d)$ cases. Consequently, $|\Phi_{sum}^{-1}(s)| = |\Phi_{sum}^{-1}(d)| + sum(l_2, -2, l_3)$ ■

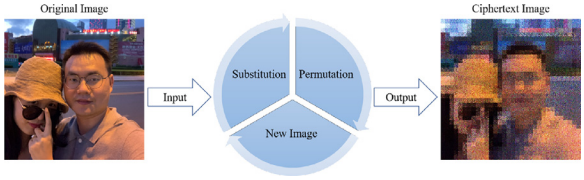

Fig. 3. The main construction of the TPE2 scheme.

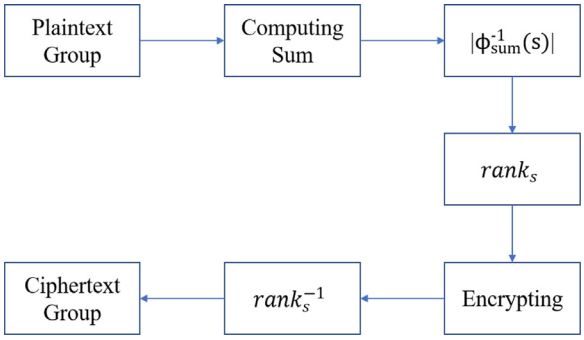

Fig. 4. The substitution encryption framework for the pixel group in the TPE2 scheme.

The main construction of the TPE2 scheme is shown in Fig. 3. It can be seen from the figure that the TPE2 scheme is mainly divided into two steps:

- **Substitution:** All pixels in the thumbnail block are divided into pixel groups (three pixels). The RTE-SPE scheme is utilized to encrypt each pixel group, which is shown in Fig. 4. Each substitution encryption has a different nonce, such as $T_{sub} = T||k||i||j||sub||num_g$ (num_g is the ordinal number of the pixel group in the block).
- **Permutation:** After all groups in the thumbnail block have completed substitution encryption, permutation encryption is carried out for all pixels in the block. Specifically, we sample a permutation from $\left\{1, \ldots, b^2\right\}$, called ω, and then shuffle the pixels in the block according to ω. Each permutation encryption has a different nonce $T_{per} = T||k||i||j||per$.

The above two steps are performed alternately on all the thumbnail blocks in the image for a complete encryption, which is called a encryption round. It is easy to know that after a complete encryption of the image, the sum of the pixels of the thumbnail blocks does not change. Meanwhile, each step of the above encryption is completely reversible, in other words, it can be decrypted lossless.

5.3. Security

In order to analyze the security of the TPE2 scheme, the encryption algorithm of the TPE2 can be modeled as Markov chain. The security of the scheme is related to the mixing time of Markov chain. The mixing time of the Markov chain refers to the minimum numbers of rounds (that is, $t_{mix}(\epsilon)$) required for the state of Markov chain to reach the ϵ-close stationary distribution [42].

First, to make it easier to understand, several definitions of Markov chain are introduced:

Definition 3. A finite Markov chain is the process of moving elements in a finite set Ω based on a transition probability matrix P.

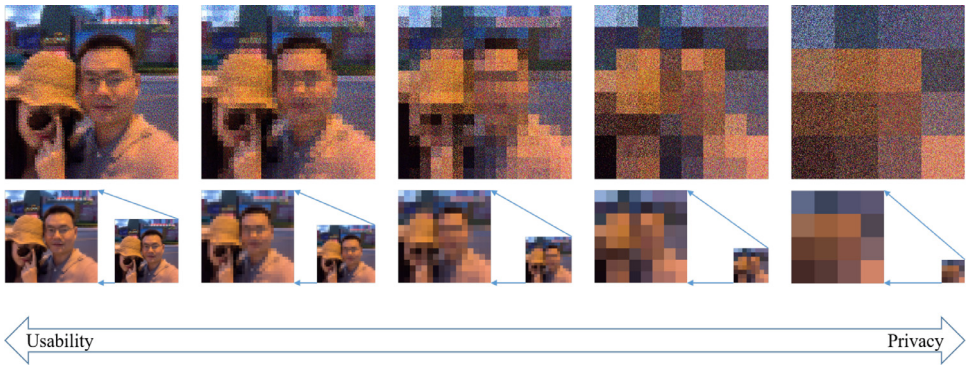

Fig. 5. The TPE2-encrypted with different block sizes (8×8, 16×16, 32×32, 64×64, and 128×128) and corresponding preserved thumbnails.

The sum of each row in P is 1 and all elements are nonnegative, that is P is stochastic [42].

Definition 4. The stationary distribution π of Markov chain is a distribution π on Ω, which makes $\pi = \pi P$ [42]. After enough rounds, the distribution of the Markov chain states approaches the stationary distribution.

Second, the TPE2 scheme is modeled as a Markov chain. For each thumbnail block, enumerate vectors in $(\mathbb{Z}_d)^n$ that have the same sum of pixel values. Each vector is a state of the Markov chain. The transition probability matrix P of the Markov chain is the probability of moving from one state to another though one encryption round. The Markov chain models the probability in the ideal manner. In other words, every substitution of a pixel group is completely uniformly chosen and the permutation (i.e., ω) of pixels in the block is also perfectly uniformly chosen.

Tajik et al. [19] proved that the uniform distribution on the set of vectors is the unique stationary distribution for the Markov chain of the TPE scheme and the chain converges to the distribution. This conclusion is also applicable to the TPE2 scheme.

Definition 5. Let our non-reversible Markov chain have transition probabilistic matrix P with $|\Omega|$ states, λ_* is the second-largest eigenvalue of the corresponding M and ϵ. The mixing time of the Markov chain can be calculated as following [19]:

$$t_{mix}(\epsilon) = \lceil \frac{2(\log \epsilon - \log(|\Omega| - 1))}{\log \lambda_*} \rceil$$

where $M = P\bar{P}$, $\bar{P} = D^{-1}PD$, $D = diag\{\pi_1, \ldots, \pi_{|\Omega|}\}$.

Tajik et al. [19] also proved that the TPE scheme satisfies NR-security when the number of encryption rounds achieves $t_{mix}(\epsilon)$. The number of encryption rounds for the TPE2 scheme to achieve NR-security also conforms to the formula. In addition, the proposed HF-TPE scheme can control the balance between privacy and usability of the ciphertext image by adjusting the dimension of the thumbnail block, as shown in Fig. 5.

6. Empirical evaluation

In this section, we report on the empirical evaluation of the TPE2 scheme and the evaluation items are listed below. *a*) The average size expansion, which is the ratio of the size of ciphertext and decrypted images to the size of original images. *b*) The perception quality of ciphertext images. *c*) The relationship between the dimension of the thumbnail block and the ability to resist face detection algorithms. *d*) The usability of ciphertext images.

Datesets: Three image datasets are utilized to evaluate the TPE2 scheme. The first dataset comprises the 500 images from the first part of the Helen dataset[2], which are transformed into 512×512 PNG images. The second dataset is 80 mugshots with the resolution of 512×512. The third dataset is 50 portraits including 4 of Jay Chou and 46 of others. In addition, some animal images and portraits of Chinese pop stars are also applied for the evaluation.

6.1. Size expansion rate

In this section, the first dataset is applied to evaluate the size expansion rate of ciphertext and decrypted images. In this evaluation, the TPE2 scheme is compared with the TPE scheme proposed by Wright et al. [14] and the TPE-LSB scheme proposed by Marohn et al. [18]. The TPE-LSB scheme approximates the original image thumbnail by adjusting the LSB of the ciphertext image. LSB(i bit) indicates that no more than i bit of the bit-plane are adjusted. To ensure the quality of the decrypted image, the maximum value of i is set to 3.

The TPE2 scheme does not generate additional data and the ciphertext image is a two-dimensional matrix as large as the dimension of the original image and therefore they should theoretically be the same size. However, image formats such as PNG make use of the information redundancy in the image to losslessly compress the image. The redundancy of the original image is often larger than that of ciphertext one, which leads to the compression ratio of the ciphertext image is lower than that of original one and thus there will be some ciphertext size expansion.

Fig. 6 (a) shows the average ciphertext image size expansion rate as thumbnail block dimensions change over original images on the first dataset. It shows that the size of images encrypted by the TPE2 scheme is about twice as large as original images. The Paillier cryptosystem may have 256 times the ciphertext expansion rate [43] and homologous encryption may even encrypt 4MB of data to 73TB [44]. In contrast, the ciphertext expansion rate of the TPE2 scheme is reasonable. Meanwhile, it can be seen from the figure that with the increase of the dimension of the thumbnail block, the ciphertext size expansion rate of each TPE scheme increases. This is mainly because the larger the block dimension, the worse the spatial locality of the pixels in the block, the compression ratio is lower. In addition, the TPE2 scheme has the highest ciphertext expansion rate since this scheme applies the substitution encryption to pixel groups, which greatly destroys the spatial locality of the image compared with other TPE schemes and this phenomenon is especially obvious when the dimension of the thumbnail block is small. The ciphertext expansion rate of the TPE2 scheme should be similar to that of Wright's scheme when the dimension of the block is large enough since the main factor that destroys the spatial locality at this time is pixel permuta-

[2] http://www.ifp.illinois.edu/vuongle2/helen/

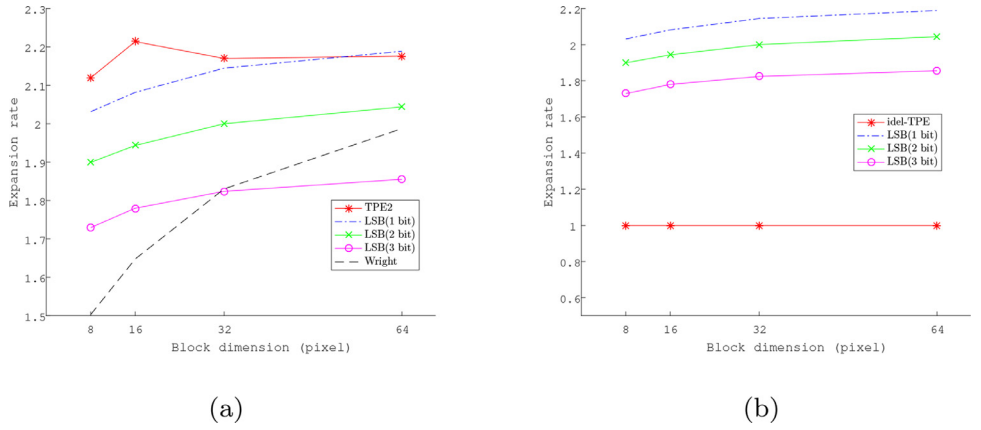

Fig. 6. 6(a) Average size expansion rate of encrypted images versus original images. 6(b) Average size expansion rate of decrypted images versus original images.

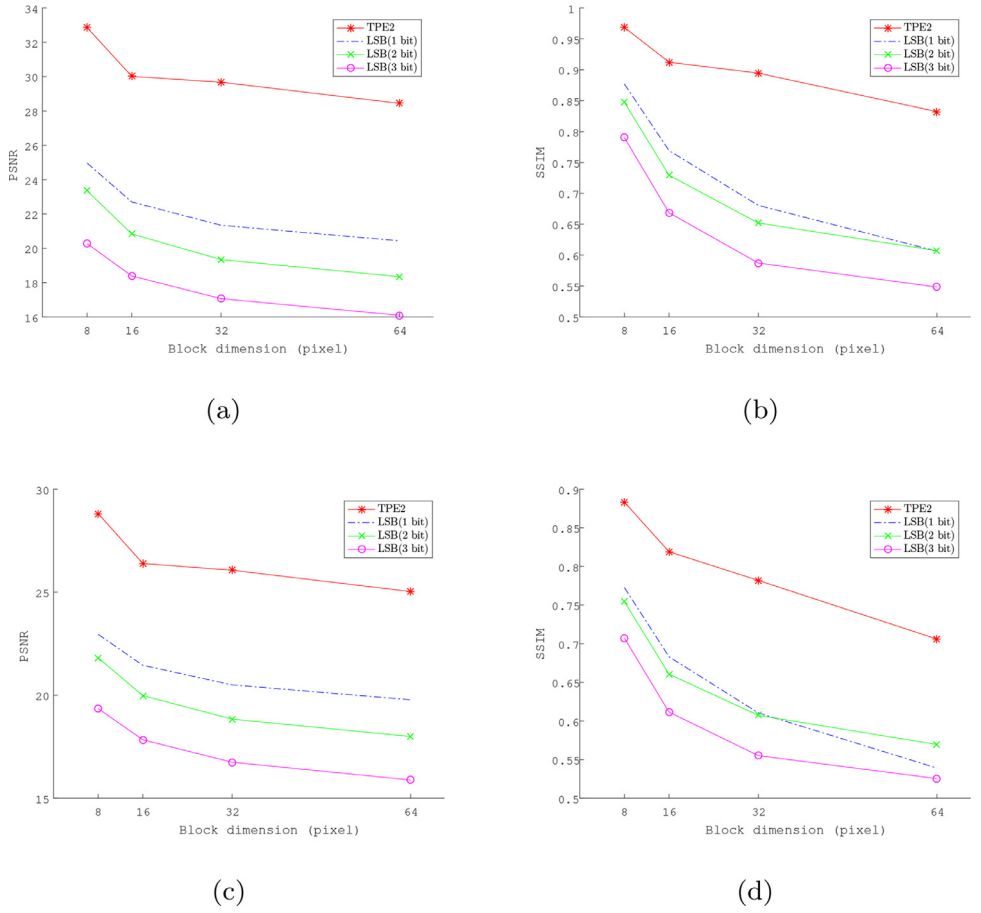

Fig. 7. 7(a) Average PSNR when thumbnail dimension is 100. 7(b) Average SSIM when thumbnail dimension is 100. 7(c) Average PSNR when thumbnail dimension is 150. 7(d) Average SSIM when thumbnail dimension is 150.

tion. The TPE-LSB scheme approximates the original image thumbnail by flipping the bit plane of the pixel, the more bit planes are flipped to a certain extent, the higher the compression ratio.

Fig. 6 (b) shows the average size expansion rate of the decrypted image as thumbnail dimensions change compared to the original image and ideal-TPE refers to the TPE scheme in which the decrypted image is the same as the original image, e.g., [14] and the TPE2 scheme. It shows that the size of the decrypted image in this scheme is far smaller than that in the TPE-LSB scheme. The size of the decrypted image is more important than the size of the ciphertext image for most users, since a critical reason for outsourcing the image to the cloud is the lack of resources of personal devices. Therefore, the TPE2 scheme is more user-friendly.

6.2. Ciphertext image perception quality

In this section, these two commonly used objective metrics for image quality assessment are applied, Peak Signal to Noise Ratio (PSNR) and Structural Similarity (SSIM), to evaluate the quality of encrypted images in the TPE2 and TPE-LSB scheme. This section does not compare other exact scheme because their perception quality is similar to that of the TPE2 scheme. In addition, thumbnail dimensions are different in various applications and systems

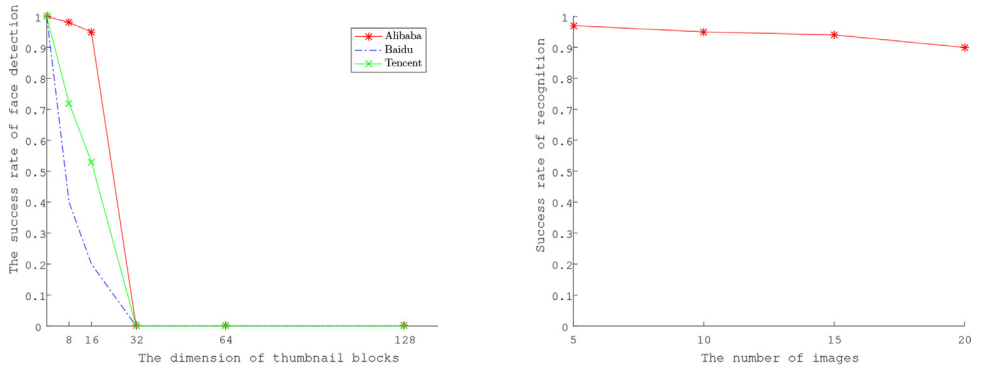

Fig. 8. 8(a) Average detection success rate of mugshots on three face detection platforms. 8(b) Respondents' average success rate in recognizing Jay Chou.

and thus two commonly used thumbnail dimension parameters are selected in the evaluation, 100×100 and 150×150. The first image dataset is utilized in this evaluation. First, the images are encrypted by the TPE2 scheme and TPE-LSB scheme, respectively. Second, the pixel values in the thumbnail block of the original image are averaged and assigned to all the pixels in the block. Third, the *imresize* function in MATLAB is utilized to adjust the ciphertext images and the images in the second step to the same thumbnail dimension (i.e., 100 and 150) and then the *psnr* and *ssim* function in MATLAB are exploited to calculate the PSNR and SSIM of the ciphertext image thumbnails, respectively.

Fig. 7 (a) and (b) show the average PSNR and SSIM of the ciphertext images compared with the original image thumbnails, respectively, when the thumbnail dimension is 100. Meanwhile, Fig. 7(c) and (d) show the average PSNR and SSIM, respectively, when the thumbnail dimension is 150. These figures demonstrate that the TPE2 scheme outperforms the TPE-LSB scheme on these two indicators. This experimentally demonstrates that the visual quality of ciphertext images in the TPE2 scheme, i.e., close to the original thumbnail, is much higher than that in the TPE-LSB scheme.

6.3. Security against facial detection

With the rapid development of computer vision technology, a large number of technologies have emerged to detect objects in images, which makes image privacy greatly threatened. Among them, face detection technology has attract great attention and is one of the most mature object technology so far. Therefore, this evaluation assessed the effect of the dimension of the thumbnail block on the success rate of face detection. In this evaluation, our TPE2 scheme encrypts the mugshots in the second image dataset with different dimensions of thumbnail blocks (as shown in Fig. 9), and then we exploit the face detection API of three state-of-the-art image analysis platforms in China, Ali Cloud Vision Platform[3], Baidu Brain[4], and Tencent AI Laboratory[5], to detect the faces in ciphertext images.

As shown in Fig. 8(a), when the dimension of the thumbnail is no less than 32, the success rate of the three face detection platforms decreases to 0. In fact, faces often make up only a small fraction of the images taken in everyday life. The following sections the dimension of the thumbnail block is 32 unless otherwise noted.

[3] https://vision.aliyun.com/
[4] https://ai.baidu.com/tech/face/
[5] https://ai.qq.com/

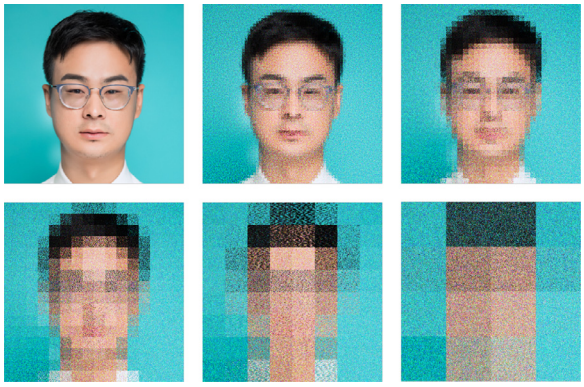

Fig. 9. TPE2-encrypted mugshots with different block sizes (plaintext, 8×8, 16×16, 32×32, 64×64, and 128×128).

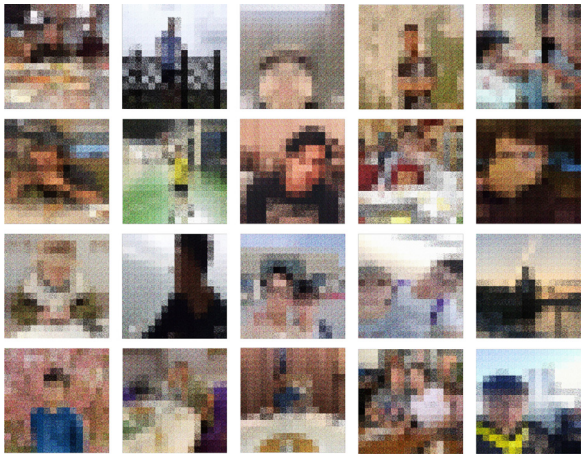

Fig. 10. 20 images encrypted by the TPE2 scheme with the dimension of the block is 32×32.

6.4. Usability evaluation

In this section, we choose images that are well known to the public in order to make the evaluation fair and objective. This allows all respondents to browse the same image sets.

6.4.1. Identifying a portrait given a name

In this section, we evaluate the recognition ability of legitimate users for encrypted images when the dimension of the block is 32×32. We choose Jay Chou's image as the target image since al-

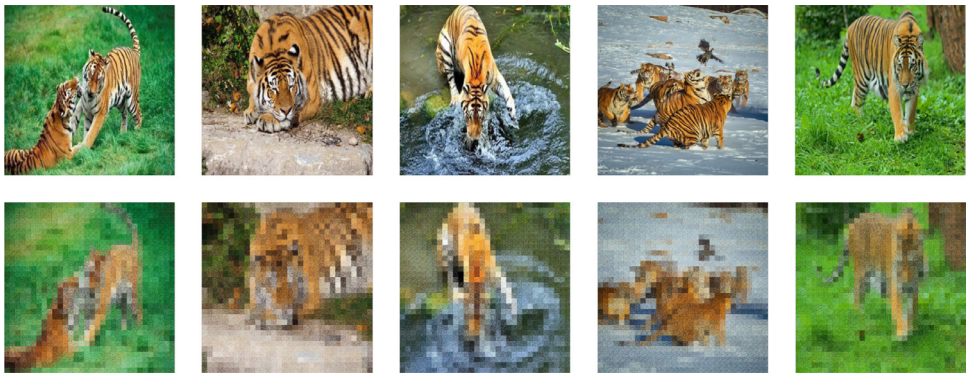

Fig. 11. Five sample images, conforming to MSD_{5H} requirements, encrypted by the TPE2 scheme with the dimension of the block is 32×32, along with their corresponding original versions. They are described as "Fighting", "Sleeping", "Dabbling", "Birding", and "Walking".

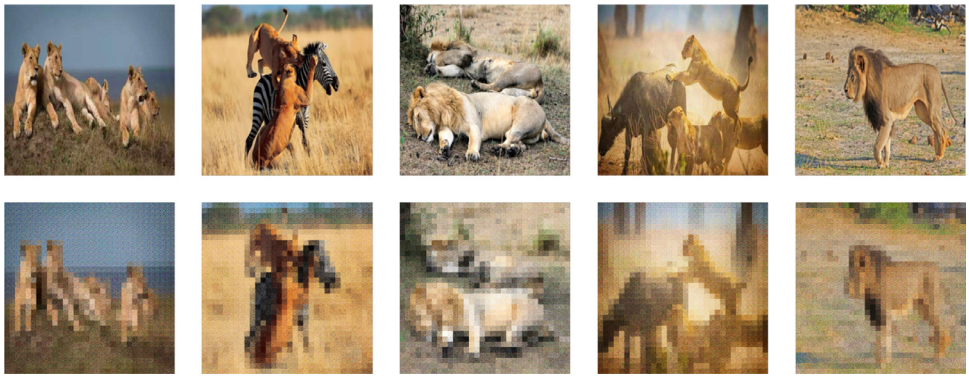

Fig. 12. Five sample images, conforming to MSD_{5L} requirements, encrypted by the TPE2 scheme with the dimension of the block is 32×32, along with their corresponding original versions. They are described as "Lying down", "Hunting, zebra", "Sleeping", "Hunting, cattle", and "Walking."

most all young people have seen Jay Chou's image, which means that there will be Jay Chou's prior knowledge, while exploiting other images may result in some respondents having no prior knowledge of the target image. To simulate users' habit of browsing albums, we divided the images in the third image set into four sets, and each image is in but only one set. These four sets have 5, 10, 15, and 20 images, respectively, and each set has one image of Jay Chou as the target image. We investigated 100 respondents and all of whom are confirmed to know Jay Chou in advance. They are asked which image in each encrypted set of images is Jay Chou, as shown in Fig. 10. To ensure that the results are not affected by the images, the 50 images are randomly grouped during every investigation.

Fig. 8 (b) shows the recognition success rate of the respondents in each set. It indicates that people who know the image are still able to accurately recognize ciphertext images when the block dimension is 32.

6.4.2. Matching scenes with descriptions (MSD)

In this section, we exploit animal images as evaluative images to test whether users are able to match the image descriptions to the corresponding encrypted images. The evaluation is divided into four parts as follow:

- MSD_{5H} : Respondents are asked to match five images with five descriptions, in which there are clear distinction between objects and backgrounds, as shown in Fig. 11.
- MSD_{5L} : Respondents are asked to match five images with five descriptions, in which the distinction between objects and backgrounds are blurred, as shown in Fig. 12.
- MSD_{10H} : This part is similar to MSD_{5H}, but 10 images and 10 descriptions are used.

Table 3
The means and standard deviations of scores for different parts of MSD.

Part	Score means	Standard deviations
MSD_{5H}	0.94	0.20
MSD_{5L}	0.91	0.23
MSD_{10H}	0.85	0.28
MSD_{10L}	0.83	0.30

- MSD_{10L} : This part is similar to MSD_{5L}, but 10 images and 10 descriptions are used.

Eighty respondents are invited to participate in the evaluation. In the part with only 5 images, the image and description matching success score is 0.2. Similarly, in the part with 10 images, the image and description matching success score is 0.1. Each part of the evaluation is required to be completed with 100 s since users usually browse quickly when choosing images in the cloud. The evaluation results are shown in Table 3. According to the evaluation, users can match image descriptions to encrypted images of which 5 images performed better than 10 because of the relatively sufficient time. If there is no experiment time is limit, the evaluation will improve a little. In addition, almost all respondents get full marks, when using the original images for the same evaluation.

6.4.3. Portrait character recognition (PCR)

In this section, we do not give the respondents any information except the images and then ask them to speak out the contents of the images. For the objective of the evaluation, we select portraits of 10 stars currently known to young Chinese people and ask respondents to name them. Each correct name is scored 0.1. We first ask 80 respondents to browse the encrypted images, then ask them

Table 4
The means and standard deviations of scores for different parts of PCR.

Part	Score means	Standard deviations
$PCR-ori$	0.85	0.13
$PCR-enc$	0.80	0.20

to browse the corresponding original ones. We record each subject's score and then calculated the means and standard deviations. The evaluation results are shown in Table 4, where $PCR-ori$ denotes the evaluation result of the respondents browsing the original images, and $PCR-enc$ denotes evaluation result of respondents browsing the encrypted ones. The results of the evaluation show that the images encrypted by the TPE2 scheme still preserve the users' ability to recognize familiar people.

7. Conclusion and prospects

In this paper, we propose the first three-pixel exact TPE scheme. This scheme provides the privacy protection and usability coexistence scheme for privacy-sensitive users who want to store local images to the cloud server for various reasons while to protect image privacy. The TPE2 scheme not only protects the privacy in images, making it impossible for the illegal third party to learn privacy information from ciphertext images, but also allows users to obtain usability from them. This allows users to safely upload images to untrusted clouds without having to choose between image privacy and usability.

Another problem that remains open is how to determine the formulations of the rank_s and rank_s^{-1} functions. In other words, the corresponding SN can be calculated directly according to the vector, or the corresponding vector can be directly calculated according to the sum of the vector elements and the SN.

Declaration of Competing Interest

No conflict of interest exits in the submission of this manuscript, and manuscript is approved by all authors for publication. Authors declare that the work described was original research that has not been published previously, and not under consideration for publication elsewhere, in whole or in part. All the authors have approved the manuscript that is enclosed.

CRediT authorship contribution statement

Ruoyu Zhao: Conceptualization, Methodology, Validation, Writing - original draft. **Yushu Zhang:** Investigation, Supervision. **Xiangli Xiao:** Formal analysis, Writing - review & editing. **Xi Ye:** Formal analysis, Writing - review & editing. **Rushi Lan:** Writing - review & editing.

Acknowledgment

The work was supported by Open Fund Project of Graduate Innovation Base (Laboratory) of Nanjing University of Aeronautics and Astronautics (No. kfjj20201601), National Natural Science Foundation of China (No. 62072237), and Guangxi Key Laboratory of Trusted Software (No. KX202027).

References

[1] M. Llc, How many photos will be taken in 2020?, 2020, https://focus.mylio.com/tech-today/how-many-photos-will-be-taken-in-2020.
[2] M. Milian, Digital photos can reveal your location, raise privacy fears, 2010, http://edition.cnn.com/2010/TECH/web/10/15/photo.gps.privacy/index.html/.
[3] M. MIRANDA, Are photographs sensitive personal information?, 2019, https://www.gmanetwork.com/news/opinion/content/697826/are-photographs-sensitive-personal-information/story/.
[4] H. Weisbaum, Digital pictures typically contain a lot of personal information, 2019, https://komonews.com/news/consumer/digital-pictures-typically-contain-a-lot-of-personal-information.
[5] A. Li, D. Darling, Q. Li, PhotoSafer: content-based and context-aware private photo protection for smartphones, in: IEEE Symp. Priv.-Aware Comput., PAC, 2018, pp. 10–18.
[6] S. Cohen, 1 second - internet live stats, 2020, https://www.internetlivestats.com/one-second/.
[7] W. Enck, P. Gilbert, B.-G. Chun, L.P. Cox, J. Jung, P. McDaniel, A.N. Sheth, TaintDroid: an information-flow tracking system for realtime privacy monitoring on smartphones, in: Proc. USENIX Conf. Operat. Syst. Des. Implement, 2010, pp. 393–407.
[8] D. Pegg, C. Cadwalladr, US data firm admits employee approached Cambridge analytica: Palantir confirm employee 'engaged in a personal capacity' with the company, 2018, https://www.theguardian.com/uk-news/2018/mar/28/palantir-employee-cambridge-analytica.
[9] Z. Schiffer, The big facebook outage offers a behind-the-scenes look at how the social network's ai "sees" your photos and interprets them for blind users, 2019, https://www.businessinsider.com/facebooks-photo-outage-reveals-how-ai-sees-your-photos-2019-7.
[10] R. Shaikh, Apple was aware about icloud flaw 6 months before celebrity photo leak, 2014, https://wccftech.com/celebrity-photo-leak-icloud/.
[11] T. Chuman, W. Sirichotedumrong, H. Kiya, Encryption-then-compression systems using grayscale-based image encryption for JPEG images, IEEE Trans. Inf. Forensic Secur. 14 (6) (2019) 1515–1525.
[12] G. Ye, C. Pan, Y. Dong, Y. Shi, X. Huang, Image encryption and hiding algorithm based on compressive sensing and random numbers insertion, Signal Process. (2020), doi:10.1016/j.optlastec.2020.106489.
[13] R. Lan, J. He, S. Wang, T. Gu, X. Luo, Integrated chaotic systems for image encryption, Signal Process. 147 (2018) 133–145.
[14] C.V. Wright, W.-c. Feng, F. Liu, Thumbnail-preserving encryption for JPEG, in: IH MMSec - Proc. ACM Workshop Inf. Hiding Multimedia Secur., 2015, pp. 141–146.
[15] A. Jolfaei, X. Wu, V. Muthukkumarasamy, On the security of permutation-only image encryption schemes, IEEE Trans. Inf. Forensic Secur. 11 (2) (2016) 235–246.
[16] S. Li, C. Li, G. Chen, N.G. Bourbakis, K.-T. Lo, A general quantitative cryptanalysis of permutation-only multimedia ciphers against plaintext attacks, Signal Process. 23 (3) (2008) 212–223.
[17] C. Li, K.-T. Lo, Optimal quantitative cryptanalysis of permutation-only multimedia ciphers against plaintext attacks, Signal Process. 91 (4) (2011) 949–954.
[18] B. Marohn, C.V. Wright, W.-c. Feng, M. Rosulek, R.B. Bobba, Approximate thumbnail preserving encryption, in: MPS - Proc. Multimed. Priv. Secur., 2017, pp. 33–43.
[19] K. Tajik, A. Gunasekaran, R. Dutta, B. Ellis, R.B. Bobba, M. Rosulek, C.V. Wright, W. Feng, Balancing image privacy and usability with thumbnail-preserving encryption, in: Proc. Symp. Netw. Distrib. Syst. Secur., 2019.
[20] M. Bellare, T. Ristenpart, P. Rogaway, T. Stegers, Format-preserving encryption, in: Lect. Notes Comput. Sci., 2009, pp. 295–312.
[21] Y. Tian, Y. Hou, J. Yuan, CAPIA: cloud assisted privacy-preserving image annotation, in: IEEE Conf. Commun. Netw. Secur., CNS, 2017, pp. 1–9.
[22] F. Markatopoulou, V. Mezaris, I. Patras, Implicit and explicit concept relations in deep neural networks for multi-label video/image annotation, IEEE Trans. Circuits Syst. Video Technol. 29 (6) (2019) 1631–1644.
[23] Y. Niu, Z. Lu, J. Wen, T. Xiang, S. Chang, Multi-modal multi-scale deep learning for large-scale image annotation, IEEE Trans. Image Process. 28 (4) (2019) 1720–1731.
[24] V. Oguz Yazici, A. Gonzalez-Garcia, A. Ramisa, B. Twardowski, J. van de Weijer, Orderless recurrent models for multi-label classification, in: IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recogn., 2020, pp. 13437–13446.
[25] D. Tao, J. Cheng, X. Gao, X. Li, C. Deng, Robust sparse coding for mobile image labeling on the cloud, IEEE Trans. Circuits Syst. Video Technol. 27 (1) (2017) 62–72.
[26] D. Tao, L. Jin, W. Liu, X. Li, Hessian regularized support vector machines for mobile image annotation on the cloud, IEEE Trans. Multimed. 15 (4) (2013) 833–844.
[27] Q. Wang, M. He, M. Du, S.S.M. Chow, R.W.F. Lai, Q. Zou, Searchable encryption over feature-rich data, IEEE Trans. Dependable Secur. Comput. 15 (3) (2018) 496–510.
[28] L. Zhang, T. Jung, C. Liu, X. Ding, X. Li, Y. Liu, Pop: Privacy-preserving outsourced photo sharing and searching for mobile devices, in: Proc. - IEEE Int. Conf. Distrib. Comput. Syst., ICDCS, 2015, pp. 308–317.
[29] Z. Xia, X. Wang, L. Zhang, Z. Qin, X. Sun, K. Ren, A privacy-preserving and copy-deterrence content-based image retrieval scheme in cloud computing, IEEE Trans. Inf. Forensic Secur. 11 (11) (2016) 2594–2608.
[30] D.G. Lowe, Distinctive image features from scale-invariant keypoints, Int. J. Comput. Vis. 60 (2) (2004) 91–110.
[31] H. Bay, A. Ess, T. Tuytelaars, L. Van Gool, Speeded-up robust features (surf), Comput. Vis. Image Underst. 110 (3) (2008) 346–359.
[32] R. Gregory, Knowledge in perception and illusion, Philos. Trans. R. Soc. Lond. B. 352 (1358) (1997) 1121–1127.
[33] K.J.G. Snodgrass, Does the generation effect occur for pictures? Am. J. Psychol. 113 (1) (2000) 95–121.

[34] G.A. Rousselet, S.J. Thorpe, M. Fabre-Thorpe, Processing of one, two or four natural scenes in humans: the limits of parallelism, Vis. Res. 44 (9) (2004) 877–894.
[35] T. Denning, K. Bowers, M. van Dijk, A. Juels, Exploring implicit memory for painless password recovery, in: Conf. Hum. Fact. Comput. Syst. Proc., 2011, pp. 2615–2618.
[36] E. von Zezschwitz, S. Ebbinghaus, H. Hussmann, A. De Luca, You can't watch this! privacy-respectful photo browsing on smartphones, in: Conf. Hum. Fact. Comput. Syst. Proc., 2016, pp. 4320–4324.
[37] L. Du, H. Ling, Preservative license plate de-identification for privacy protection, in: Proc. Int. Conf. Doc. Anal. Recognit., 2011, pp. 468–472.
[38] T. Orekondy, M. Fritz, B. Schiele, Connecting pixels to privacy and utility: automatic redaction of private information in images, in: IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recogn., 2018, pp. 8466–8475.
[39] E.Y. Sun, H. Wu, C. Busch, S.C. Huang, Y. Kuan, S.Y. Chang, Efficient recoverable cryptographic mosaic technique by permutations, IEEE Trans. Circuits Syst. Video Technol. (2020), doi:10.1109/TCSVT.2020.2976050.
[40] J. He, B. Liu, D. Kong, X. Bao, N. Wang, H. Jin, G. Kesidis, PUPPIES: transformation-supported personalized privacy preserving partial image sharing, in: IEEE/IFIP Int. Conf. Dependable Syst. Networks, DSN, 2016, pp. 359–370.
[41] H. wen Xue, J. Du, S. liang Li, W. jiao Ma, Region of interest encryption for color images based on a hyperchaotic system with three positive Lyapunov exponets, Opt. Laser. Technol. 106 (2018) 506–516.
[42] D. Levin, Y. Peres, Markov chains and mixing times, Am. Math. Soc., 2006.
[43] P. Paillier, Public-key cryptosystems based on composite degree residuosity classes, in: J. Stern (Ed.), Lect. Notes Comput. Sci., 1999, pp. 223–238.
[44] Z. Brakerski, C. Gentry, V. Vaikuntanathan, (Leveled) fully homomorphic encryption without bootstrapping, in: Commun. Comput. Info. Sci., 2012, pp. 309–325.