IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received February 17, 2020, accepted March 18, 2020, date of publication March 25, 2020, date of current version April 10, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2983194

An AES-Based Secure Image Retrieval Scheme Using Random Mapping and BOW in Cloud Computing

HUA WANG[1,2], ZHIHUA XIA[1,2], (Member, IEEE), JIANWEI FEI[1,2], AND FENGJUN XIAO[3]

[1] Engineering Research Center of Digital Forensics, Ministry of Education, Nanjing University of Information Science and Technology, Nanjing 210044, China
[2] Jiangsu Engineering Center of Network Monitoring, Jiangsu Collaborative Innovation Center on Atmospheric Environment and Equipment Technology, School of Computer and Software, Nanjing University of Information Science and Technology, Nanjing 210044, China
[3] Management School, Hangzhou Dianzi University, Hangzhou 310088, China

Corresponding author: Zhihua Xia (xia_zhihua@163.com)

This work was supported in part by the Jiangsu Basic Research Programs-Natural Science Foundation under grant numbers BK20181407, in part by the National Natural Science Foundation of China under grant numbers U1936118, 61672294, in part by Six peak talent project of Jiangsu Province (R2016L13), Qing Lan Project of Jiangsu Province, and "333" project of Jiangsu Province, in part by the National Natural Science Foundation of China under grant numbers U1836208, 61702276, 61772283, 61602253, and 61601236, in part by National Key R&D Program of China under grant 2018YFB1003205, in part by the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAPD) fund, in part by the Collaborative Innovation Center of Atmospheric Environment and Equipment Technology (CICAEET) fund, China. Zhihua Xia is supported by BK21+ program from the Ministry of Education of Korea.

ABSTRACT With the rapid growth of the number of images, many content-based image retrieval methods have been extensively used in our daily life. In general, image retrieval services are very expensive in terms of computing and storage. Therefore, outsourcing services to the cloud server is a good choice for image owners. However, privacy protection can become a big issue for image owners because the cloud server can only be semi-trusted. In this paper, we propose a novel image retrieval scheme. It is a ciphertext image retrieval method based on random mapping features with the bag-of-words model. After encrypting the image with Advanced Encryption Standard and block permutation, the cloud server generates random templates and then extracts the local features. All local features are clustered by k-means algorithm to form the visual word. The histogram of encrypted visual words is constructed in this way as the feature vector to represent each image. The similarity between images can be measured by the distance between feature vectors on the cloud server. Experiments and analysis prove the effect of the scheme.

INDEX TERMS Image retrieval, AES encryption, BOW model, random mapping.

I. INTRODUCTION

With the rapid development of imaging technology, various image information has been widely used in our daily life and the number of images has increased dramatically. To achieve the efficient use of the image information, the researchers proposed some Content-based Image Retrieval (CBIR) schemes to improve the efficiency of retrieving similar images [1]. However, a typical image database is usually very large, including millions of images. It becomes difficult for image owners to store and search images efficiently with such a large number of images. Therefore, CBIR services typically require large amounts of storage and computation. These requirements make it attractive to outsource CBIR services to the cloud server because the cloud server usually has more storage space and stronger computing power than the image owner. In this way, the image owner does not need to store the image database locally. People can retrieve images directly in the cloud server.

In addition to the huge benefits of CBIR outsourcing, image privacy is a top concern for the image owner. Images may contain all kinds of sensitive information, such as the current location, educational background, family members, and personal interests. Besides, on the one hand, hackers are usually interested in the information stored on the cloud server, so the cloud server can often be attacked by hackers. In 2017, Yahoo confirmed that 3 billion user accounts of its cloud server had been affected by hacker attacks. The stolen

The associate editor coordinating the review of this manuscript and approving it for publication was Akansha Singh.

information includes user name, email address, telephone number, birthday and some other account information. On the other hand, the cloud server can correctly follow some specified protocol specifications, but can be interested in inferring or analyzing information from the image to learn sensitive information. Therefore, it is not safe to directly outsource CBIR to a semi-honest cloud server. The cloud server may analyze the private information of images, thereby revealing personal privacy. To protect the privacy of the images, people can encrypt the image. However, the encrypted image may not be suitable for image retrieval in this way.

At present, researchers have proposed many privacy-preserving image retrieval schemes, which can mainly be divided into two categories. The first category is the feature-encryption based image retrieval scheme. First, the image owner extracts some features from the plaintext image. Then, all plaintext images and plaintext features are encrypted and uploaded to the cloud server, respectively. This kind of scheme needs to ensure that the similarity between images can be measured by comparing the distance of encrypted features. The other category is the image-encryption based image retrieval scheme. In this kind of scheme, the image owner uses some suitable encryption methods to encrypt the image, so as to ensure that the effective features can be extracted from the encrypted image for image retrieval. After that, all encrypted images will be uploaded to the cloud server. The task of feature extraction and image retrieval is outsourced to the cloud server to reduce the computing and storage burden of the image owner. In general, the second kind of scheme is more practical because it outsources most computing and storage tasks to the cloud server. However, there are still some shortcomings in the existing schemes, such as image security, retrieval accuracy and so on.

Contribution. We have proposed an AES-based image retrieval scheme using random mapping and the bag-of-words (BOW) model. The contributions can be summarized as follows:

1) An image encryption method is proposed using the Advanced Encryption Standard (AES) encryption and block permutation for the image searchable scheme. The proposed encryption method is secure under the known-plaintext attack, which is securer than many symmetric searchable schemes.
2) A feature extraction method is proposed to calculate useful features from AES-encrypted images using random mapping and the BOW model. The proposed method can successfully retrieve similar images to some extends. It could be an interesting discovery to the field of searchable encryption, or even to the field of cryptography.

The rest of the paper is arranged in the following order. Section II introduces some related works of image retrieval. Section III tells the technology we used and presents an overview of the scheme. Section IV describes the design of the proposed scheme in detail. Section V presents the security analysis of the proposed scheme. The experiment results and the retrieval efficiency of our scheme are shown in section VI. At last, a summary is drawn in Section VII.

II. RELATED WORK

Since CBIR has been proposed, through the continuous deepening and expansion of researchers, the achievements of these studies are well applied in work and life. CBIR aims to find similar images. The similarity between images is judged by the distance of the features extracted from the images. However, plaintext images cannot be outsourced to the cloud directly due to privacy issues. It is worth noting that, in addition to the original image data, the features may also leak information about the image content. Specifically, the ciphertext image retrieval is mainly divided into two categories, one is the feature-encryption based image retrieval schemes, and the other is the image-encryption based image retrieval schemes.

In the feature-encryption based image retrieval schemes, the image owner extracts features from the plaintext image. Then, all images and their corresponding features are encrypted by the image owner. Lu *et al.* [2] proposed three feature protection methods with cryptographic techniques, including bit-plane randomization, random projection, and randomized unary encoding. The bit-plane randomization and randomized unary coding support the calculation of the Hamming distance between the encrypted features. The random projection supports the approximate calculation of L1 distance. In another scheme, Lu *et al.* [3] proposed two schemes for implementing secure retrieval in the ciphertext domain, namely the secure inverted index and the Min-Hash algorithm. The order-preserving encryption and the random hash function were used to encrypt the occurrence frequency of image features. The Jaccard distance was used to measure the similarity between images. Lu *et al.* [4] compared their previous encryption methods in [2] with the homomorphic encryption method. All the methods above will decrease retrieval accuracy while the homomorphic encryption can achieve the same retrieval accuracy with that in plaintext domain. However, the homomorphic encryption scheme is at a disadvantage in terms of retrieval efficiency and computational cost as it involves additional communication, storage, and computation. Xia *et al.* [5] proposed a ciphertext image retrieval scheme based on matrix transformation. The image features were protected by multiplying them with the inverse matrix. The similarity of images was compared by calculating the Euclidean distance of features. The scheme of Jiawei *et al.* [6] used secure KNN to perform a secure image retrieval, and additionally built a tree index to improve image retrieval efficiency. Wang *et al.* [7] used the secure modular hash to extract features for image retrieval. The hash bits are clustered by the k-means algorithm to improve the efficiency of image retrieval. Weng *et al.* [8], [9] proposed two privacy-preserving schemes and obtained the hash value of the image by locally sensitive hash (LSH). They ignore or encrypt some hash bits to hide the image information for security. The cloud server uses the remaining hash bits to perform the similarity search and return similar images to the user. Finally, the returned images are optimized by the hash bits at the user side to obtain similar images. Zhang *et al.* [10] extracted features by using the homomorphic encryption algorithm, and used attribute-based encryption to enhance image security. Finally, the cloud server returned similar images by comparing features. Xia *et al.* [11] proposed to extract features based on the SIFT feature and the BOW model. The Earth Mover's Distance (EMD) is used to calculate the distance between image features. To protect parameter information, they do a linear transformation on EMD. Besides, LSH is used to improve retrieval efficiency. Xia *et al.* [12], [13] extracted features to represent corresponding images and used the secure KNN algorithm to protect image features. They used LSH to improve retrieval efficiency and introduced an encryption-domain watermarking method to prevent the propagation of illegal images. Qin *et al.* [14] proposed a scheme with extracting features by the Harris algorithm. After that, they used LSH to construct the index and improve retrieva efficiency.

In the above feature-encryption based image retrieval schemes, the image owner extracts features from the plaintext image. The image owner encrypts images and features to realize secure image retrieval, which takes a lot of computational overhead. Therefore, it is desirable to outsource most tasks to the cloud server. In image-encryption based image retrieval schemes, the image owner only encrypts the images while the feature extraction and the search operation are performed on the cloud server. Some scholars have studied the special image encryption algorithm and the extraction of effective features from the encrypted image to achieve ciphertext image retrieval. The key is how to extract effective features from encrypted images. Abduljabbar *et al.* [15] proposed a ciphertext image retrieval scheme. They used AES encryption to encrypt the plaintext image and extracted the local SURF features from the encrypted image. Then, they compared the similarity of images by calculating the Euclidean distance and improved retrieval efficiency by LSH. Cheng *et al.* [16] proposed a retrieval scheme based on the Markov model. They used stream cipher encryption and scrambling encryption to encrypt JPEG images, and then extracted Markov features related to discrete cosine transform (DCT) coefficients as features. Besides, Cheng *et al.* [17] proposed an encrypted JPEG image retrieval scheme that extracts features based on local AC coefficients. The DC coefficients and the quantization tables are encrypted by the stream cipher encryption. The intra-block positions and inter-block positions of DCT coefficient blocks in the same component are shuffled. They counted the histogram of all the AC coefficients and calculated the similarity of the image by comparing the distance between the blocks, which is time-consuming. Cheng *et al.* [18] proposed another encrypted JPEG image retrieval scheme, which extracted features based on local variances. Image content was protected by stream cipher encryption and permutation encryption. Gong *et al.* [19] proposed a scheme based on orthogonal decomposition. They built a framework that splits the feature vector into two parts. One part is encrypted for image privacy and the other is used to extract image features. Ferreira *et al.* [20] proposed a new ciphertext retrieval scheme. They used value permutation, pixel position scrambling to protect the values and positions of pixels. The histogram of encrypted color value is calculated as images features. The similarity between images is measured by the Hamming distance.

In this paper, we propose an outsourced CBIR scheme based on the BOW model, AES encryption, block permutation and random mapping. The proposed scheme outsources the tasks of feature extraction, index construction and search operations to the cloud server.

III. SYSTEM OVERVIEW

A. SYSTEM MODEL

The proposed scheme involves two different entities, namely image owner and cloud server, as shown in Figure 1.

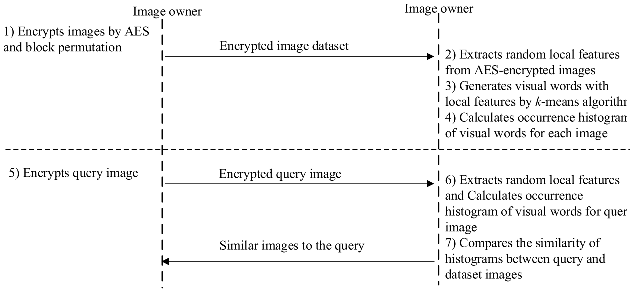


Image owner
Image owner
1) Encrypts images by AES and block permutation
Encrypted image dataset
2) Extracts random local features from AES-encrypted images
3) Generates visual words with local features by *k*-means algorithr
4) Calculates occurrence histogran of visual words for each image
5) Encrypts query image
Encrypted query image
6) Extracts random local features and Calculates occurrence histogram of visual words for quer image
Similar images to the query
7) Compares the similarity of histograms between query and dataset images

FIGURE 1. **A system model of the proposed scheme.**

The image owner has a large image database $I\left\{I_i\right\}_{i=1}^{n}$, with a corresponding identity set $ID\left\{ID_i\right\}_{i=1}^{n}$. To protect privacy information, the image database needs to be encrypted to generate an encrypted image set $C\left\{C_i\right\}_{i=1}^{n}$. In addition to image encryption, the image owner wants to outsource most computing and storage tasks to the cloud server whenever possible. Besides, the image owner wants to use the CBIR service provided by the cloud server to efficiently retrieve images similar to the query image. However, for the privacy of the query image, the image should be encrypted at first.

In our scheme, the feature extraction and search operations are performed by the cloud server. The cloud server stores the encrypted image database for the image owner and provides the image retrieval service for the image owner. In our scheme, the cloud server is responsible for generating the index in addition to computing and searching for similar images.

B. SECURITY MODEL

Similar to previous methods, we first assume that the cloud server is semi-honest. The cloud server can honestly follow the specified rules, but will be curious to analyze the data of sensitive information. The image owner is trustworthy, and will not leak any information to the cloud server. We do not consider information leakage due to the access pattern.

C. PRELIMINARIES

1) BAG-OF-WORDS MODEL

The purpose of feature extraction is to reduce storage requirements, retrieval time and improve retrieval accuracy. Global features can be extracted from the entire image, while local features are extracted from small regions. In general, local features are more robust and may have better retrieval effects. An effective method of local features is the BOW model. The BOW model first appeared in the field of text information retrieval. This model treats documents as collections of words, ignoring connections between words. After that, the BOW model was widely used in image retrieval and got good results. In this model, local features are extracted from the images. They cluster together, and the cluster center is defined as a visual word to form the vocabulary. Then, all local features are represented by their most recent visual words and the image is represented as a histogram of visual words. The BOW model has three steps, which is described as follows:

Local feature extraction. Local features will be extracted from all images at first. Suppose that each image in the database composed of image blocks, and each image block can be a feature vector. Feature vectors can be selected according to the specific way. Some common features include color histogram, SIFT, LBP, etc.

Vocabulary construction. Then, a corresponding vocabulary should be built. The process of constructing the vocabulary needs to cooperate with some clustering algorithms, and the k-means algorithm is used here. The clustering center of the algorithm is defined as a visual word used for the operation of clustering all the extracted local features. Then, the vocabulary can be composed of visual words.

Histogram calculation. Finally, the histogram of visual words should be calculated. All local features will be represented by visual words. The corresponding histograms of visual words generated by images are used to measure the similarity between images.

Based on the BOW model of the plaintext domain, the bag-of-encrypted-words model in the ciphertext domain is constructed. The steps of the model are as follows:

The first step is the extraction of local features. Features extracted from encrypted image blocks can be used as local features of the image. The second step is to create an encrypted vocabulary. Using the k-means algorithm, the cluster center can be an encrypted visual word of the encrypted vocabulary. The third step is to construct a histogram of the word frequency. The local feature in the image will be approximated represented by the encrypted word in the encrypted vocabulary. Then, the word frequency of the encrypted word is counted together to construct an encrypted word frequency histogram.

2) AES ENCRYPTION

AES, also known as Rijndael encryption in cryptography, is a block encryption standard adopted by the USA government. AES encryption groups the plaintext content and each block is equal in length. It encrypts one data block at a time until the entire plaintext encrypted. In the AES encryption, the length of a plaintext block is fixed to 128 bits. Each byte is 8 bits, that is, each block is 16 bytes. However, there are several options for the length of the key, such as 128 bits, 192 bits or 256 bits. AES encryption does not only encrypt once, but it also requires multiple rounds of encryption. The number of encryption rounds is different for different keys.

In our scheme, the length of the encryption key is 128 bits and the encryption process needs to be performed for ten rounds. The encryption steps performed in the first nine rounds are the same. Each round performs four operations in order: byte substitution, row shift, column mixing and round key addition. Besides, column mixing is not necessary for the last round. It is worth mentioning that the plaintext block and the initial key matrix need to be XOR before the first round of processing. The 128-bit key can be changed to a 4×4-byte matrix, and each column of the matrix is treated as a word, which constitutes a key sequence in order. Next, we will extend the length of the key sequence array. The key expansion involves three steps: word loop, byte substitution and round constant XOR. AES decryption is also divided into ten rounds, and the decryption process is the reverse process of the encryption process.

IV. THE PROPOSED SCHEME

A. OVERVIEW OF THE SCHEME

The proposed scheme consists of two entities, namely the image owner and the cloud server. The two entities have their own tasks. The image owner is responsible for the key generation and image encryption. The cloud server executes the index generation and the search. The image owner and the cloud server run the trapdoor generation together. At last, the image owner is responsible for the image decryption.

The image owner owns the image database $I = \{I_i\}_{i=1}^{n}$ and the corresponding identity set $ID\{ID_i\}_{i=1}^{n}$. It runs the key generation algorithm to generate the secret key $\mathcal{K}$, and the secret keys are stored in the image owner. The image owner then runs the image encryption algorithm to encrypt the image database, generate an encrypted image database $C\{C_i\}_{i=1}^{n}$ and upload it to the cloud server. After receiving the encrypted image database, the cloud server runs an index generation algorithm to build the index.

To perform the retrieval, the owner runs the trapdoor generation algorithm to generate trapdoors and submits a query request to the cloud server to complete the remaining steps. The cloud server runs the search algorithm to retrieve similar images, and these similar images are returned to the image owner as search results. After receiving the search results, the image owner runs the decryption algorithm to decrypt images.

The summary of the algorithms is shown in Algorithm 1.

Algorithm 1 Overview of the Algorithms

Image owner:

- The key generation algorithm takes the security parameters and returns the secret keys.
- The image encryption algorithm takes the secret keys $\mathcal{K}$, the image database I, the identity set ID, and the block size $Bblksize$ as inputs, and returns the encrypted image database.
- The trapdoor generation algorithm takes the secret keys $\mathcal{K}$, the query image, and the block size $Bblksize$ as inputs, and returns the trapdoor.
- The image decryption algorithm takes the secret keys $\mathcal{K}$, the encrypted similar image set, and the block size $Bblksize$ as inputs, and returns the decrypted image set.

Cloud server:

- The index generation algorithm takes encrypted image database C and the block size $Bblksize$ as inputs, and returns the index.
- The search algorithm takes the encrypted database C, the index, the trapdoor, and the block size $Bblksize$ as inputs, and returns encrypted similar images.

B. IMAGE ENCRYPTION

Next, we will introduce the specific encryption process of our scheme in detail. The entire process of image encryption is defined as follows.

1) KEY GENERATION

In the proposed scheme, the images are encrypted by AES and block permutation. This scheme handles images in the RGB color space, and the three channels are encrypted by AES with different keys key_R, key_G, and key_B. In addition, images are further encrypted by block permutation with different keys for different images. Thus, we need a pseudorandom permutation generator (RPG) and a master key (key_M) to generate random permutations. Summarily, the secret keys of our scheme include

$$\mathcal{K} = \{key_R, key_G, key_B, key_M, RPG\}, \tag{1}$$

where key_R, key_G, key_B are 128-bit sequences. AES encrypts the images by 4×4 image blocks. And the block permutation is conducted on big-blocks which consist of several adjacent 4×4 image blocks and denoted as $Bblk$. We generate a distinct permutation for each as

$$pmt \leftarrow \mathrm{RPG}(key_M, ID, Bblknum), \tag{2}$$

where ID is the identity of the corresponding image, $Bblknum$ is the number of big-blocks, and pmt is a random permutation of $\{1, \ldots, Bblknum\}$.

2) AES-BASED ENCRYPTION

AES is a typical block cipher. Here we choose the block size of 128 bits. For the three color channels of the image I, we divided each channel, $I_i, i \in R, G, B$, into 4×4 blocks which are rightly 128 bits. These blocks are encrypted by AES separately. The process of AES-based encryption is summarized in Algorithm 2.

Algorithm 2 AES-Based Encryption

```
Input: Original image $I$ and secret keys $key_i$, $i \in R, G, B$
Output: AES-encrypted image $I'$
For $\forall\ I_i, i \in R, G, B$ do
    Divide the image $I_i$ into $4 \times 4$ non-overlapping
    image blocks denoted as $blk$
    For $\forall\ blk_j \in I_i$ do
        $blk'_j \leftarrow \mathrm{AES}(blk_j, key_i)$
        Put the encrypted block $blk'_j$ in the
        corresponding position of $I'_i$;
    End for
End for
```

3) BLOCK PERMUTATION

After AES encryption, we further disturb the image content by permutation. Here, the permutation is also conducted by blocks. We named these blocks as big-blocks, denoted as $Bblk$, which can consist of one or more adjacent 4×4 blocks. For an image, the three channels are permuted with the same key to facilitate the following feature extraction. For different images, the permutation key can be different for better security. The process of block permutation is summarized in Algorithm 3.

C. FEATURE EXTRACTION

In our scheme, the feature extraction is also outsourced to the cloud server, reducing the computational burden on the image

Algorithm 3 Block Permutation

```
Input: An AES-encrypted image $I'$ with three channels
$I'_i, i \in R, G, B$, image identity $ID$, and number of big-
blocks $Bblknum$
Output: Encrypted image $C$
For $\forall I'_i, i \in R, G, B$ do
    Divide $I'_i$ into big-blocks denoted as $Bblk_j, j = 0,$
     $\ldots, Bblknum$, where $Bblknum$ is the number of
     big-blocks;
    Generate a permutation for the image as $pmt$
     $\leftarrow \mathrm{RPG}(key_M, ID, Bblknum)$;
    For $\forall Bblk_j, j = 1, \ldots, Bblknum$ do
     $Bblk'_j \leftarrow Bblk'_{pmt[j]}$;
     Put the encrypted block $Bblk'_j$ in the corresponding
     position of $C_i$;
    End for
End for
```

owner. The feature extraction of images mainly consists of three steps.

1) LOCAL FEATURE EXTRACTION

Generally, the AES-encrypted image can be regarded as totally random. It is hard to know how to extract meaningful features from such data. Thus, we try to extract features randomly. We divide images into big-blocks, which are made up of several adjacent 4×4 blocks and denoted as $Bblk$. Then, local feature vectors are extracted from such big-blocks.

Firstly, we generate θ templates $T_j, j = 1, \ldots, \theta$ with the size of $(4n) \times (4m)$. The elements of the template are the random integer values in a fixed range $[-\tau, \tau]$. Then, the images are also divided into big-blocks with the size of $(4n) \times (4m)$. Next, the result production of a big-block and each of the templates is calculated as

$$p_j = Bblk \cdot T_j. \tag{3}$$

where the value of the image big-block and the corresponding value of the random template are multiplied and then added, as shown in Fig.2. Then, p_j is shrunk by arctan function to avoid the values with extremely absolute as

$$p_j \leftarrow \arctan(p_j). \tag{4}$$

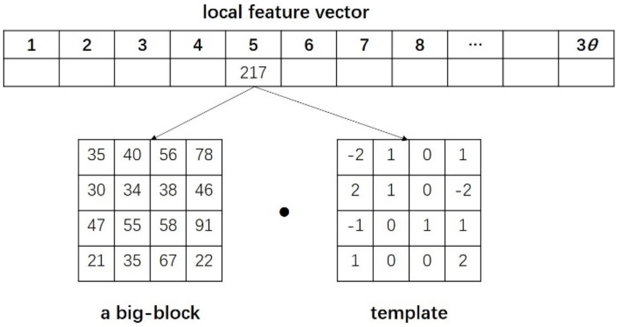


FIGURE 2. An example of calculation of local feature.

The p_j can be calculated from the three channels of the image. In this way, a local feature vector $\mathbf{p} = (p_1, \ldots, p_{3\theta})$ is calculated to represent a corresponding big-block.

2) VOCABULARY GENERATION

The k-means algorithm is used to generate the vocabulary. Firstly, a set of local feature vectors $\{\mathbf{p}_{i,j}\}$ are calculated from all of the images in the dataset, with $\mathbf{p}_{i,j}$ denotes the j-th local feature vector of i-th image. Then, the local features are clustered into k classes by using the k-means algorithm. The cluster centers are defined as the words to form the vocabulary. It is worth noting that the visual word is also a θ-dimensional vector.

3) GLOBAL FEATURE CALCULATION

Here, we generate a global feature vector from an image to represent it. For an image, we can find the closest visual word to each local feature vector. Then, occurrence histograms of the visual words are calculated and normalized, generating a k-dimensional feature vector $\mathbf{f} = (f_i)_{i=1}^{k}$. Then, the similarity between the images can be measured by the distance of their feature vectors.

D. IMAGE RETRIEVAL

1) TRAPDOOR GENERATION

In our scheme, the trapdoor generation is mainly divided into the following steps. First, the image owner encrypts the image and uploads it to the cloud. After receiving the encrypted query image, the cloud server divides it into big-blocks and calculates the set of local feature vectors $\{\mathbf{p}_{q,j}\}_{j=1}^{Bblknum}$. Finally, the global feature vector $\mathbf{q} = (q_i)_{i=1}^{k}$ is calculated as the trapdoor.

2) SEARCH OPERATION

The similarity between the query feature vector and the feature vector of a dataset image are measured by Manhattan distance as

$$d(\mathbf{q}, \mathbf{f}) = \sum\nolimits_{i}^{k} |q_i - f_i| \,. \tag{5}$$

where $\mathbf{q}$ and $\mathbf{f}$ denote the feature vector of the query image and a dataset image. Finally, the images with the smallest distance are returned as search results.

3) IMAGE DECRYPTION

The owner needs to decrypt these images after receiving encrypted images. The decryption process is completely opposite to the encryption process.

V. SECURITY ANALYSIS

This paper considers an honest-but-curious cloud server, which honestly provides the storage and search service to the image owner, but is curious to the data of the image owner. In this section, we analyze the information leakage and the security of our scheme.

A. INFORMATION LEAKAGE

The proposed scheme encrypts the image by AES and block permutation. The leaked information includes the sizes of images and big-blocks. To find similar images efficiently, the similarity between images is leaked to the cloud server. That is to say, the cloud server knows which images are similar to each other. We named this leakage as the similarity pattern. In addition, the cloud server knows which images are returned to the image owner in every search. Thus, the access pattern is also leaked to the cloud server. The leakage of the similarity pattern and the access pattern are the common trade-offs to efficient search in searchable encryption schemes.

B. SECURITY OF IMAGE CONTENT

With the leaked information above, the cloud server cannot recover the image content yet. Firstly, the image pixels are encrypted by AES, which is secure under the known-plaintext

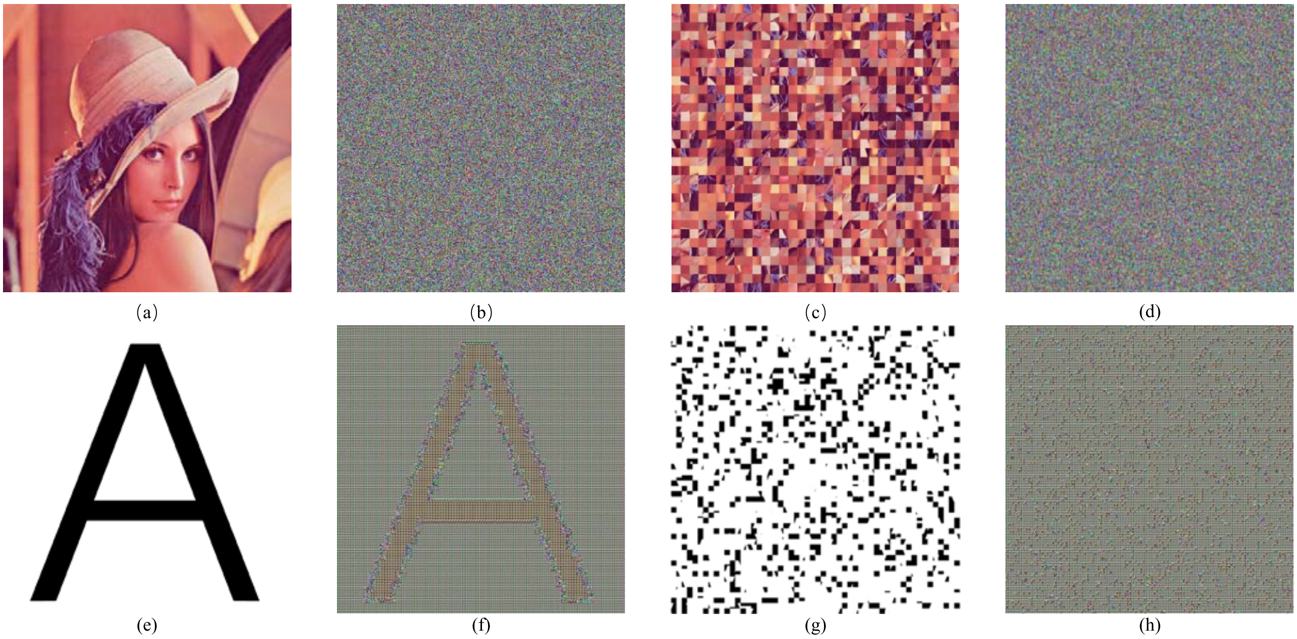


FIGURE 3. Visual effects of encryption: (a) and (e) are original images; (b) and (f) are encrypted images with AES; (c) and (g) are encrypted images with block permutation; (d) and (h) are encrypted images with AES and block permutation.

attack. It has better security than many other symmetric searchable schemes that are only secure under ciphertext-only attack. Besides, we permute image blocks to protect the image content. We generate a distinct secret key for each image, which means one-time pad encryption and cannot be cracked.

C. SECURITY OF IMAGE FEATURE

In this paper, the image features are extracted randomly from the encrypted images. The cloud server can learn nothing from such features. Besides, the cloud server can only extract features from the images encrypted by the correct secret key for image search. That is to say, the cloud server cannot generate effective feature without owner-encrypted images.

VI. EXPERIMENTAL RESULTS

This section tests the performance of the proposed scheme. A prototype system is implemented with Matlab 2016a and runs on a Win10 system with an Intel Core (TM) i7-7700HQ (CPU 2.80 GHz) and 16 GB memory. The Inria Holidays database is used here as the experimental database. It includes 1491 color images in total, 500 queries and 991 corresponding relevant images, the size of which is about 3264 × 2448. The Inria Holidays database provides a Python evaluation package to calculate mAP and is used in many image retrieval schemes, which facilitates fair retrieval accuracy comparison.

A. EFFECTIVENESS OF ENCRYPTION

In the proposed scheme, images are encrypted by AES and block permutation. AES is a worldwide used encryption method and can provide very good security. As shown in Figure 3(b), the content of a nature image is disrupted. However, if the image content is quite simple, the encrypted image will still expose some shape information as shown in Figure 3(f). The block permutation can further improve security. As shown in Figure 3(d) and (h), the images can be well protected by the combination of AES and block permutation.

B. RETRIEVAL ACCURACY

In the test of retrieval accuracy, the Python evaluation package provided by the Inria Holidays database was used to calculate the mean average precision (mAP). The retrieval accuracy of the proposed scheme mainly depends on several parameters, namely the number of random templates θ, the size of big-blocks $Bblksize$, and the number of cluster centers k, as listed in Table 1.

TABLE 1. Symbol of parameters.

Parameters	Symbol
Size of big-blocks	$Bblksize$
Number of random templates	θ
Number of cluster centers	k

We first test the retrieval accuracy of the proposed scheme with different numbers of random templates θ and cluster centers k. As shown in Table 2, the proposed scheme achieves the best retrieval accuracy with $\theta = 1000$ and $k = 1000$.

Next, we test the retrieval accuracy of the proposed scheme under different sizes size of big-blocks with $\theta = 1000$

TABLE 2. Retrieval accuracies with different θ and k.

	k				
θ	100	200	500	1000	2000
100	0.2821	0.2834	0.2853	0.2861	0.2851
200	0.2829	0.2838	0.2857	0.2865	0.2854
500	0.2841	0.2852	0.2858	0.2864	0.2856
1000	0.2848	0.2860	0.2866	**0.2872**	0.2863
2000	0.2835	0.2845	0.2853	0.2862	0.2855

and $k = 1000$. As shown in Table 3, the proposed scheme achieves best with $Bblksize = 4 \times 4$.

TABLE 3. Retrieval accuracies with different *Bblksize*.

Bblksize	mAP
4×4	0.2872
16×16	0.2516
64×64	0.2214
256×256	0.1937

The experiments show that the proposed scheme can retrieve similar images successfully to some extent. However, the retrieval accuracy is not so good when compared with previous schemes, as shown in Table 4. However, the encryption methods in previous schemes are only secure under the ciphertext-only attack while the proposed scheme is secure under the known-plaintext domain. Besides, we demonstrate that useful features can be extracted from AES-encrypted images. It is an interesting discovery.

TABLE 4. mAP of different schemes.

Schemes	mAP
Lu[3]	0.49
Cheng[16]	0.54
Ferreira[20]	0.55
Xia[21]	0.67
Ours	0.28

C. TIME CONSUMPTIONS

1) TIME CONSUMPTION OF IMAGE ENCRYPTION

In the proposed scheme, the images are encrypted by AES and block permutation. The time consumption of AES is fixed while the time consumption of block permutation depends on the size of the big-block. Table 5 lists the time consumption of different *Bblksize* in the encryption of the whole Inria Holidays database.

2) TIME CONSUMPTION OF FEATURE EXTRACTION

In the proposed scheme, the feature extraction includes three steps, local feature extraction, vocabulary generation, and global feature calculation.

In the extraction of local features, the time consumption depends on the number of cluster centers k and the size of big-blocks. Table 6 lists the time consumption of local feature extraction from the whole Inria Holidays database with different k and *Bblksize*.

TABLE 5. Time consumption of encryption (s).

Bblksize	4×4	16×16	64×64	256×256
Time consumption of AES	1842126.21			
Time consumption of block permutation	1434.02	390.13	284.36	254.32

TABLE 6. Time consumption of local feature extraction (s).

	k				
Bblksize	100	200	500	1000	2000
4×4	78555	135599	303615	642680	1056003
16×16	11817	21919	52437	13035	210680
64×64	4207	7965	19114	51702	93517
256×256	3692	7035	17039	42644	85469

TABLE 7. Time consumption of vocabulary generation (s).

	k				
Bblksize	100	200	500	1000	2000
4×4	4179.65	6862.64	12843.64	17004.63	19794.81
16×16	1652.06	2511.13	4109.81	4337.63	5754.75
64×64	970.78	1436.75	2211.25	2613.28	3497.96
256×256	311.24	436.11	519.35	671.71	724.71

TABLE 8. Time consumption of global feature calculation (s).

	k				
Bblksize	100	200	500	1000	2000
4×4	691.46	1142.57	2108.79	3851.65	7068.79
16×16	204.79	287.52	546.21	945,36	1637.16
64×64	18.01	29.80	43.89	81.52	145.68
256×256	8.96	9.75	11.38	14.35	18.49

TABLE 9. Time consumption of search (s).

k	100	200	500	1000	2000
Time consumption	0.0062	0.0086	0.0148	0.0202	0.0358

In the proposed scheme, the vocabulary is constructed by the k-means cluster algorithm. The time consumption depends on the number of cluster centers k and the size of big-blocks *Bblksize*, as shown in Table 7. Please note that the number of big-blocks *Bblknum* is negatively associated with the size of big-blocks *Bblksize*.

The time consumption of global feature calculation depends on the number of cluster centers k and the size of big-blocks *Bblksize*. Table 8 lists the time consumption of local feature extraction from the whole Inria Holidays database with different k and *Bblksize*.

3) TIME CONSUMPTION OF LINEAR SEARCH

A linear index is used in our scheme so the cloud server needs to search the entire index to find the most similar images. Besides, the number of cluster centers k determines the dimension of the feature vector, thereby affecting the time consumption of search. The time complexity of this traversal search is $O(k \times n)$. Table 9 presents the time consumption of the search with different k.

VII. CONCLUSION

This paper proposes a new privacy-preserving image retrieval scheme. The images are encrypted by AES and block permutation. The features are extracted from AES-encrypted images by random mapping and bag-of-words. The encrypted images are secure under the known-plaintext attack. The experiments demonstrate the proposed scheme can retrieval similar images to some extent. Generally, it is hard to believe one can extract useful features from AES-encrypted images for similarity comparison. However, our experiments demonstrated that the randomly extracted features are useful for searching for similar images. This could be an interesting discovery. In the future, it is possible to extract better local features from AES-encrypted image blocks. On the other hand, AES may needs improvements to prevent the extraction of meaningful features.

ACKNOWLEDGMENT

The authors would like to thank the Editor and the Anonymous Reviewers for their constructive comments and suggestions, which improve the quality of this article.

REFERENCES

[1] Y. Liu, D. Zhang, G. Lu, and W.-Y. Ma, "A survey of content-based image retrieval with high-level semantics," *Pattern Recognit.*, vol. 40, no. 1, pp. 262–282, Jan. 2007.

[2] W. Lu, A. L. Varna, A. Swaminathan, and M. Wu, "Secure image retrieval through feature protection," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Apr. 2009, pp. 1533–1536.

[3] W. Lu, A. Swaminathan, A. L. Varna, and M. Wu, "Enabling search over encrypted multimedia databases," *Proc. SPIE*, vol. 7254, Feb. 2009, Art. no. 725418.

[4] W. Lu, A. L. Varna, and M. Wu, "Confidentiality-preserving image search: A comparative study between homomorphic encryption and distance-preserving randomization," *IEEE Access*, vol. 2, pp. 125–141, 2014.

[5] Z. Xia, Y. Zhu, X. Sun, and J. Wang, "A similarity search scheme over encrypted cloud images based on secure transformation," *Int. J. Future Gener. Commun. Netw.*, vol. 6, no. 6, pp. 71–80, Dec. 2013.

[6] J. Yuan, S. Yu, and L. Guo, "SEISA: Secure and efficient encrypted image search with access control," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, Apr. 2015, pp. 2083–2091.

[7] Y. Wang, M. Miao, J. Shen, and J. Wang, "Towards efficient privacy-preserving encrypted image search in cloud computing," *Soft Comput.*, vol. 23, no. 6, pp. 2101–2112, Mar. 2019.

[8] L. Weng, L. Amsaleg, A. Morton, and S. Marchand-Maillet, "A privacy-preserving framework for large-scale content-based information retrieval," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 1, pp. 152–167, Jan. 2015.

[9] L. Weng, L. Amsaleg, and T. Furon, "Privacy-preserving outsourced media search," *IEEE Trans. Knowl. Data Eng.*, vol. 28, no. 10, pp. 2738–2751, Oct. 2016.

[10] L. Zhang, T. Jung, K. Liu, X.-Y. Li, X. Ding, J. Gu, and Y. Liu, "PIC: Enable large-scale privacy preserving content-based image search on cloud," *IEEE Trans. Parallel Distrib. Syst.*, vol. 28, no. 11, pp. 3258–3271, Nov. 2017.

[11] Z. Xia, Y. Zhu, X. Sun, Z. Qin, and K. Ren, "Towards privacy-preserving content-based image retrieval in cloud computing," *IEEE Trans. Cloud Comput.*, vol. 6, no. 1, pp. 276–286, Jan. 2018.

[12] Z. Xia, X. Wang, L. Zhang, Z. Qin, X. Sun, and K. Ren, "A privacy-preserving and copy-deterrence content-based image retrieval scheme in cloud computing," *IEEE Trans. Inf. Forensics Security*, vol. 11, no. 11, pp. 2594–2608, Nov. 2016.

[13] Z. Xia, N. N. Xiong, A. V. Vasilakos, and X. Sun, "EPCBIR: An efficient and privacy-preserving content-based image retrieval scheme in cloud computing," *Inf. Sci.*, vol. 387, pp. 195–204, May 2017.

[14] J. Qin, H. Li, X. Xiang, Y. Tan, W. Pan, W. Ma, and N. N. Xiong, "An encrypted image retrieval method based on Harris corner optimization and LSH in cloud computing," *IEEE Access*, vol. 7, pp. 24626–24633, 2019.

[15] Z. A. Abduljabbar, H. Jin, A. Ibrahim, Z. A. Hussien, M. A. Hussain, S. H. Abbdal, and D. Zou, "Privacy-preserving image retrieval in IoT-cloud," in *Proc. IEEE Trustcom/BigDataSE/ISPA*, Aug. 2016, pp. 799–806.

[16] H. Cheng, X. Zhang, J. Yu, and F. Li, "Markov process-based retrieval for encrypted JPEG images," *EURASIP J. Inf. Secur.*, vol. 2016, no. 1, pp. 1–9, Dec. 2016.

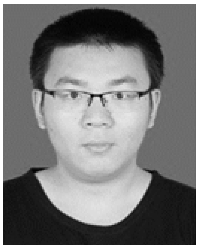
[17] H. Cheng, X. Zhang, J. Yu, and Y. Zhang, "Encrypted JPEG image retrieval using block-wise feature comparison," *J. Vis. Commun. Image Represent.*, vol. 40, pp. 111–117, Oct. 2016.

[18] H. Cheng, J. Wang, M. Wang, and S. Zhong, "Toward privacy-preserving JPEG image retrieval," *Proc. SPIE*, vol. 26, no. 4, 2017, Art. no. 043022.

[19] J. Gong, Y. Xu, and X. Zhao, "A privacy-preserving image retrieval method based on improved BoVW model in cloud environment," *IETE Tech. Rev.*, vol. 35, no. 1, pp. 76–84, Dec. 2018.

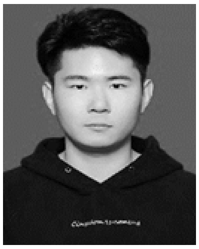
[20] B. Ferreira, J. Rodrigues, J. Leitao, and H. Domingos, "Practical privacy-preserving content-based retrieval in cloud image repositories," *IEEE Trans. Cloud Comput.*, vol. 7, no. 3, pp. 784–798, Jul. 2019, doi: 10.1109/TCC.2017.2669999.

[21] Z. Xia, L. Jiang, D. Liu, L. Lu, and B. Jeon, "BOEW: A content-based image retrieval scheme using bag-of-encrypted-words in cloud computing," *IEEE Trans. Services Comput.*, early access, Jul. 10, 2019, doi: 10.1109/TSC.2019.2927215.



HUA WANG was born in Nantong, Jiangsu, China, in 1994. He received the B.S. degree in computer science and technology from the Nanjing University of Information Science and Technology, Nanjing, Jiangsu, in 2017, where he is currently pursuing the M.S. degree. His research interests include encrypted image retrieval, cloud computing security, and multimedia processing.

ZHIHUA XIA (Member, IEEE) received the B.S. degree from Hunan City University, China, and the Ph.D. degree in computer science and technology from Hunan University, China, in 2006 and 2011, respectively. He is currently an Associate Professor with the School of Computer and Software, Nanjing University of Information Science and Technology. His research interests include digital forensic and encrypted image processing.



JIANWEI FEI was born in Nanjing, Jiangsu, China, in 1996. He received the B.E. degree in electronic and information engineering from Nanjing Forestry University, in 2018. He is currently pursuing the master's degree in computer science with the Nanjing University of Information Science and Technology. His research interests include artificial intelligence security and multimedia forensics.

FENGJUN XIAO received the B.S. degree in economics from Beihang University, in 2009, and the master's degree in technology policy, in 2014, under the supervision of Prof. Shi Li. He is currently pursuing the Ph.D. degree under the supervision of Prof. Chengzhi Li. Since 2015, he has been on Research in network security and emergency management.