SMPP-CBIR: shorted and mixed aggregated image features for privacy-preserving content-based image retrieval

Ali Lazim Lafta, Ayad I. Abdulsada
Department of Computer Science, Education College for Pure Sciences, University of Basrah, Basrah, Iraq

ABSTRACT

Thanks to recent breakthroughs in photographic and digital technology, enormous amounts of image data are generated daily. Many content-based image retrieval (CBIR) systems have been developed for searching image collections. However, these systems need more computer and storage resources that can be met by cloud servers, since they supply a lot of processing power at a reasonable price. The protection of users' personal information is a worry for image owners since cloud services are not exactly trustworthy. In this paper, we suggest and put into practice a CBIR (SMPP-CBIR) technique for searching and retrieving ciphertext information that protects security. Asymmetric scalar-product-preserving encryption process (ASPE) is used to preserve aggregated mixed feature vectors while still enabling computation between them to describe the related picture collection. The k-means clustering algorithm is used to recursively arrange all encrypted attributes into a tree index in order to speed up search times. The findings show that SMPP-CBIR is more scalable, more precise, and faster in indexing and retrieval than earlier systems.

Keywords: Aggregated image features, Content-based image retrieval, Privacy-preserve, Searchable encryption, VLAD

This is an open access article under the CC BY-SA license.

1. INTRODUCTION

Content-based image retrieval (CBIR) is a helpful approach for searching image collections and finding comparable images that has been used for many years in a variety of real-world applications such as face recognition, object identification, and medical detection. The increasing use of digital cameras and cellphones, on the other hand, has resulted in massive image archives. As a result, standard CBIR methods will be forbidden since they need more storage and computational resources. Cloud computing can assist by giving data owners with on-demand access to sufficient storage and computational resources. In this scenario, images will be outsourced to the cloud server and no longer be under the supervision of their owner [1]. CBIR may be used by authorized users to connect the cloud server and get comparable images. Because images are frequently personal and include sensitive information, sending them directly to the cloud poses a significant privacy risk. Patients' images [2], for example, are not allowed to be shared with anybody other than a specific doctor in the medical use of CBIR. Most of the time, images are encrypted before being transferred to cloud servers to decrease the danger of privacy being compromised. CBIR activities will be disabled if basic encryption techniques are used directly. As a result, developing privacy preserver CBIR (PP-CBIR) systems that can deal with encrypted images without decryption is critical.

The following is how the existing privacy-preserving CBIR schemes work: the data owner extracts certain feature vectors from the image. Then, before being sent to the cloud server, all images and vectors are
encrypted. In this case, the distance between two images' matching encrypted characteristics can be used to calculate their similarity. Image feature vectors can be global, which creates a summary vector for the entire image, or local, which represents the image by its interest spots, resulting in a large number of feature vectors. On the other hand, global features are dependent on the image's signal representation; any change in the illumination, scale, rotation, or color depth in the same image will result in a new feature vector.

To describe the image, several approaches for global features, such as shape [3], [4], color histograms [5], and texture [6], [7], can be employed. On the other hand, the local feature will be determined by the image's interest point, such as edges, angles, or tiny image patches. Rotations, scaling, color depths, and other effects can make interest spots more resistant. The interest point will not be based on a single pixel, but rather on the pixels around it. SIFT, SURF, ORB, and LBP [8]–[11], the most well-known local feature image descriptors, each with its unique length; SIFT has d=128 dimensional with positive values. In this paper, we create our shorted and mixed image features that combined half the size of the aggregations local feature descriptors VLAD [12] with half the size of the global MPEG-7 visual descriptors [4], as far we know this is the first time to be presented as shorted mix image features using the combination of local and global image features. Many PP-CBIR methods utilize homomorphic encryption (HE) to safeguard the aggregated vectors, which permits certain arithmetic operations on the encrypted data. On the other hand, HE entails a great deal of intricacy [13]. Instead, we used the ASPE approach, which was developed by Wong et al. in 2009 [11] and used by many constructions such as [14]. In the encryption domain, this approach can easily implement kNN similarity. However, to check the submitted query to the current encrypted vectors, the cloud server must do a large number of operations. To address this problem and boost search efficiency, we build a hierarchal-indexing technique that uses the k-means clustering algorithm further work will plan to add the deep learning [15] methods as clustering algorithms. The encryption key must be shared with authorized data users who create the trapdoor for their query image. The data user's privacy is protected in this option since the data owner has no idea what the user is looking for.

Our contributions

In order to create shorted and mixed image features, we combine half of the local feature descriptor aggregations VLAD [12] with half of the global feature MPEG-7 visual descriptors. This process will allow for quick searching and indexing, as well as the mixed features will hold both the strong qualities of local and global visual features. Moreover, we employ ASPE, a lightweight encryption method with superior scalability and efficiency, to protect the aggregated vectors. However, we adopt the k-means clustering method to build a hierarchal index to boost search performance. Furthermore, we combine the most well-known global descriptor MPEG-7 with two popular local descriptors.

2. RELATED WORKS

The two modes in which the existing PP-CBIR schemes operate are encrypted features schemes and encrypted image schemes. In the first method, images features are extracted and encrypted before being stored in the cloud services provider by the data owner. In the second method, images are encrypted and feature extraction in the encryption domain is delegated to the cloud server-side.

2.1. Encrypted features schemes

In the past years, much research tries to fix the problems of PP-CBIR. Lu et al. [16] suggested PP-CBIR in the encrypted domain where images are described as global histograms of visual words. Such histograms are encrypted by either order-preserving encryption or min-hash functions in order to find the similarity between two images, Jaccard distance was used to measure the distance between their histograms. Xia et al. [17] proposed to use ASPE to secure global features. To encrypt the features, they used a binary vector to split each feature vector into two vectors. Then, they defined two invertible matrices to encrypt each split feature vector. This will enable the cloud server to calculate the similarity distance between two image vectors in the encryption domain without any communication between the data owner and the cloud server. However, the authors used global features to describe images, whereas our scheme uses aggregated shorted and mixed features. Xia et al. [18] designed a PP-CBIR that will use the bag of visual words (BOVW) model to represent the image based on SIFT local features. The earth mover’s distance (EMD) is used to measure the distance between two images. EMD is calculated by constructing and solving a linear programming problem, the above-mentioned method was used to ensure that sensitive information to be protected does not leak out. However, one of its problems is that it needs to communicate more than once between the data owner and the cloud provider to find images that can be close to the query image, this situation will greatly increase the time taken to complete the search process and incurs high communication cost.
2.2. Encrypted image schemes

Cheng et al. [19] proposed a PP-CBIR scheme that works only with JPEG images. The retrieval accuracy of the scheme is further improved by [20], Ferreira et al. [21], [22] used full image cipher as every pixel in the image and hamming distance is used to measure the distance between two images. Wang et al. [23] proposed to extract random features from images that are encrypted by AES. Xia et al. [24] had encrypts images in YUV from this encrypted image, two histograms are extracted. Here, the Manhattan distance is used. Xia et al. [25] encrypt the entire image and represent it in encrypted histograms by using the BOVW model. Xia et al. [26] present a secure pixel and block shuffling are integrated to create a privacy-protected LBP extraction method in the encrypted field. Xu et al. [27] proposed PP-CBIR the image is divided into two symmetric parts, the first part is protected using the AES encryption method. As for the second part, it remains the same, it will be used to extract the image features, according to the above, this method will leak a lot of information about the content of the image.

3. METHOD

3.1. Scheme description

We divide our proposed SMP-CBIR scheme into three main entities: the first is the data owner, the second is the authorized data user, and the third is the cloud service provider as the workflow illustrated in Figure 1.

![Figure 1. Proposed scheme SMP-CBIR](image)

a. The data owner: plans to outsource his private image collection \( M = \{ m_1, m_2, ..., m_n \} \) of \( n \) images in its encrypted format \( C = (c_1, c_2, ..., c_n) \) to an external cloud server, with the aim of enabling the search over the encrypted collection. In the beginning, the data owner extracts aggregated shorted and mixed feature vectors \( V = (v_1, v_2, ..., v_n) \) from the plaintext image collection, and then create our secure index tree \( I \) from \( V \). Then both \( C \) and \( I \) will be both stored in the cloud server. The data owner should authorize the data users via a specific authentication scheme, which is outside the scope of our work as many existing PP-CBIR schemes [12], [17], [21]-[30].

b. The data users: are the users who authorized by data owner and want to search query images in the encrypted collection. To retrieve images, data user must provide a valid search trapdoor \( TR \) to the cloud server. When he/she gets the encrypted results, he/she will use the secret keys provided by the data owner to decrypt the encrypted results.

c. The cloud server: provides the responsibility to store the encrypted image collection with its encrypted index and supports computational power needed to answer the search requests of data users.
3.2. Design goals
Our goals summarized in the below points:
a. Efficiency: the search linearily is completely ineffective and impractical by default for huge image collections. Our proposed uses a secure tree index to achieve better search efficiency.
b. Data privacy: the actual content of the image collection, image features, and search requests should remain secret to the semi-trusted cloud server.

3.3. Shortened and mixed image features
The image features need to be perfected to deal with large image collections, several quantize methods (aggregate) developed to compact image feature into single descriptors vector in tradeoff with precision. vector of mix locally aggregated descriptors, will merge half the local feature descriptor f that represented by VLAD of l- dimensions, where the l = (k * d). with the global image features MPEG - 7 (F) with half the dimensions d of F as (1):

\[ v_{i,j} = \left[ \sum_{f_{j},f \in \text{Vlad}} f_j - \theta_{i,j} \right] \frac{1}{\sqrt{2}} \| F_{\text{MPEG-7}} \|_2 d \]  

where \( i = 1, ..., k, j = 1, ..., d \). Finally, \( L_2 \) normalization is applied to VLAD vector. Our system give flexible options to mix half the dimension of the local descriptors vector of locally aggregated descriptors (VLAD) [12] with half dimension of the global descriptors F of the five MPEG-7 [31] visual descriptors, thus for this, we will have a mix image feature representations with the reduce size. Therefore, our scheme will use the particularly good retrieval accuracy resulting from the local descriptors with the high efficiency for search of the global descriptors.

3.4. Proposed scheme
3.4.1. Privacy-preserving CBIR scheme
Please note that the data owner runs KeyGen and IndexGen, the data user runs TradoorGen and ImgDec, and the cloud server runs Search. In this subsection, we explain these algorithms in detail.

- a. K→KeyGen (λ) algorithm will receive the security parameter \( \lambda \) and returns the set key K=(S, M1, M2, kcoll), where is a binary vector of (l + 1) bits. M1 is an invertible matrix of size (l + 1)×(l + 1). M2 is defined in the same of M1, kcoll it is the secret key that will be used for encryption and decryption of images and image features.

- b. (C, I)→IndexGen (K, M) this algorithm takes as inputs K and M and returns the encrypted image collection C and the secure index I. Images could be encrypted using any secure method.

- c. TR→TradoorGen (K, m_q) the data user will run the TradoorGen to generate the trapdoor for his query image \( m_q \), to retrieve similar images from the cloud server and it should not leak any information to the cloud server about the query image or the results.

- d. \( \phi \)→Search (I, TR, C) when the cloud server receives the trapdoor TR from the data user, the Search algorithm will find the most similar cluster, we compare our query trapdoor against all its descriptors. Calculating the distance in the encryption domain will be as follows:

\[ v_i^T v_q^T = (\delta M_{1_i}^{-1} \hat{v}_{qa})^T M_{1_i}^{-1} \hat{v}_{ia} + (\delta M_{2_i}^{-1} \hat{v}_{qb})^T M_{2_i}^{-1} \hat{v}_{ib} \]
\[ = \delta(\hat{v}_{qa})^T \hat{v}_{ia} + \gamma(\hat{v}_{qb})^T \hat{v}_{ib} \]
\[ = \delta(\|v_q\| - 2 \sum_{j=1}^l v_{i,j} v_{q,j}) \]
\[ = \delta(\|v_q - v_i\|^2 - 2 \|v_q\|^2) \]  

The cloud server will have the ability for to find the closest feature vectors without revealing the original aggregated shorted and mix feature vectors \( v_{i,j} \). The final step is to send the top-\( \phi \) similar encrypted images to the data user.

3.5. Security analysis
In this part we will discuss the security issues of our proposed scheme.

- a. Image content privacy: images could be encrypted with any standard method for data encryption. Thus, we will not consider its security as these methods are well defined and proved. The illegal distribution of the retrieved images could be prevented using data hiding techniques [32].

- b. Mix and shorted aggregated features privacy: recall that aggregated feature vectors are protected by ASPE method [33] which is proved to be secure against ciphertext-only attacks.

SMPP-CBIR: shorted and mixed aggregated image features for privacy-preserving ... (Ali Lazim Lafta)
c. Query trapdoor privacy: the query image trapdoors are generated and encrypted by the same method for aggregated image vectors. Thus, they are all well protected too.

d. Access and search pattern: like earlier PP-CBIR schemes, our scheme leaks the access and search pattern to the cloud server. Such information can be protected but at the expense of more computation and communication costs.

4. RESULTS AND DISCUSSION

During our experiments, we used the precision metric to measure the retrieval effectiveness, which is defined as \( Pr = \emptyset / \emptyset \), \( \emptyset \) standing for the real number of the relevant images that are retrieved. Notice that, the similarity (2) will be conducted over encrypted vectors without affecting the precision. We employed two and mixed feature descriptors:

To test the retrieval precision, we given 20 image queries from the ten different categories because our experiments was done on Corel-1k image data base as showing in Figure 2. Therefore, the retrieval precisions are the average values of 20 search queries. Figures 3(a) and (b) shows the average retrieval precision for different \( \emptyset \) values. Recall that mixed and shorted aggregating image feature are generated from half the local descriptors VLAD with \( k \) visual words complained with half the global image feature MPEG-7. Our experiments are conducted for different \( k \) values: 2, 4, 8, 16, 32 as visual words. Notice that SIFT descriptors are slightly better than ORB descriptors if complained with the same MPEG-7 descriptor.

![Figure 2. The 10 categories of Corel-1k dataset](image)

4.1. Efficiency investigation

In this subsection, we investigate the efficiency of our scheme in terms of time consumption. The time consumption is presented according to the index creation, trapdoor generation and search operation.

a. Index construction time; recall that the secure index is constructed as a tree index from the mix aggregated features for the entire image collection. Figures 4(a) and (b) illustrates the index construction times for a variable number of images \( n \) with different number of \( (k) \) visual words.

b. Trapdoor generation time; Figures 5(a) and (b) reports the trapdoor generation time for different mix descriptors with different number of \( (k) \) visual words.

c. Search time; Figures 6(a) and (b) illustrates the search time for a variable number of images with different variations of aggregated mix vectors.
Figure 3. Average precision (a) half VLAD using ORB descriptors with half MPEG-7 and (b) half VLAD using SIFT descriptors with half MPEG-7.

Figure 4. The time cost of secure index construction (a) half VLAD using ORB descriptors with MPEG-7 and (b) HVLAD using SIFT descriptors with half MPEG-7.

Figure 5. The time cost of trapdoor generation (a) half VLAD using ORB descriptors with half MPEG-7 and (b) half VLAD using SIFT descriptors with half MPEG-7.

SMPP-CBIR: shorted and mixed aggregated image features for privacy-preserving ... (Ali Lazim Lafta)
5. CONCLUSION
In our paper, we designed and applied a new SMPP-CBIR scheme within the setting of the cloud computing. Each image is described as a single compact mix aggregated vector that is derived half of it from the local descriptors and the other half from the global descriptors. This method significantly reduces the computation and commination costs. The shortened and mixed aggregated feature vectors are encrypted using ASPE algorithm, which enables the cloud server to calculate the resemblance scores for the encrypted image feature vectors without decryption or any added round of communication. The shortened and mix image feature vectors are indexed as tree-index to improve the search efficiency from $O(n)$ to $O(n^2)$. Our experiments are performed in many scenarios of shortened and mix aggregated feature vectors were generated with a variable number of visual words and global descriptors. Results illustrate the practical value of our proposed scheme. For future work, we try to embed invisible watermarks for preventing dishonest users from the illegal distribution of images.

REFERENCES


BIOGRAPHIES OF AUTHORS

Ali Lazim Lafta is an MSc student in the Computer Department of the Education College for Pure Sciences at the University of Basra. In 2011, he earned his BSc in computer science from the University of Basrah's Department of Computer Science at the Education College for Pure Sciences in Iraq. His areas of interest in research include secure image retrieval, water marking, image representations and cyber security. He can be contacted at email: alialmalki000122@gmail.com.

Prof. Dr. Ayad I. Abdul-Sada is a Professor in the Computer Science Department at the University of Basrah. He Get his Ph.D degree from Huazhong University in China in Computer Science at 2013. He received his master degree from Basrah University at 2005. His areas of interest in research include searchable encryption, advanced cryptography, information outsourcing, and cloud security. He can be contacted at email: ayad.abdulsada@uobasrah.edu.iq.