Secure Image Retrieval Scheme in the Encrypted Domain

Pei Zhang[1], Li Zhuo[1], Yingdi Zhao[1], Bo Cheng[1], Jing Zhang[1] and Xiaoqin Song[2]
[1] *Signal & Information Processing Laboratory, Beijing University of Technology, Beijing, China*
[2]*College of Electronic and Information Engineering, Nanjing University of Aeronautics and Astronautics, Nanjing, China*

Keywords: Secure Image Retrieval, Feature Encryption, Encrypted Domain, Content-based Image Retrieval.

Abstract: Currently, the image retrieval methods focus on improving the retrieval performance, but ignoring preserving the problem of preserving privacy. Images contain a great deal of personal privacy information, and leakage of information will result in seriously negative effect. Ensuring the image retrieval performance while preserving the confidentiality of data has become the key issue in the field of image retrieval. Based on the Content-based Image Retrieval (CBIR), we propose a secure image retrieval scheme in the encrypted domain, where the encrypted features can be used in similarity comparison directly. This paper compares the ciphertext retrieval with plaintext retrieval to illustrate that the proposed scheme could achieve the comparable retrieval performance, while ensuring the image information security at the same time.

1 INTRODUCTION

Image retrieval is an effective technique to find the images that the users need from the vast mass image database accurately and quickly. In the past decade, most works of image retrieval focus on how to improve the retrieval performance (Wang et al., 2010); (Gao et al., 2012,). But less work takes into account the security of the image content. Malicious intruders and tampers will result in serious problems on many aspects, such as personal privacy, political and military events, as well as business transactions. Therefore, how to efficiently ensure the security of the image content information has become a key issue in the field of image retrieval.

The key problem of image information protection is the processing of secure signal To ensure the security of the image content, many encryption methods can be used, such as AES (Zeghid et al., 2007) and cryptographic primitives (Erkin et al., 2007). When these encryption methods are applied to the CBIR (Content-based Image Retrieval), the features extracted from the encrypted image may not be used directly. The reason is that the essence of the CBIR is to compare the similarity among high dimensional image features (Eakins, and Graham, 1999), but the similarity cannot be preserved after the images are encrypted by the encryption method above. In this way, features used in CBIR should be extracted from the decrypted image. However, the features of the decrypted image may leak image information. Moreover, when the image database is very large, decrypting each image will cost great computation and time for search.

Recent work by Lu et al. (Lu et al., 2009) proposed two secure indexing schemes built upon visual words representation of images by using signal processing and cryptographic techniques jointly. Both indexing schemes realize image retrieval efficiently while preserving the confidentiality of data. The work by Lu et al. (Lu et al., 2009) built three secure CBIR schemes, including bitplane randomization, random projection and randomized unary encoding, where protection of image features allowed the similarity comparison among encrypted features. And these works by Lu et al. are the first endeavors on the CBIR in the encrypted domain.

This paper presents an secure image retrieval scheme in the encrypted domain based on the CBIR framework. The proposed scheme can not only ensure the security of image information, but also achieve comparable retrieval performance with plaintext retrieval. Compared with the work of Lu et al. (Lu et al., 2009), this paper makes an improvement on the feature extraction methods to obtain better image retrieval performance in the encryption domain. Specific steps of the feature extraction and encryption are provided.

Zhang P., Zhuo L., Zhao Y., Cheng B., Zhang J. and Song X..
Secure Image Retrieval Scheme in the Encrypted Domain.
DOI: 10.5220/0004289707830787
In *Proceedings of the International Conference on Computer Vision Theory and Applications* (VISAPP-2013), pages 783-787
ISBN: 978-989-8565-47-1

2 THE IMAGE RETRIEVAL IN THE ENCRYPTED DOMAIN

The framework of the secure image retrieval in the encrypted domain is shown in Figure 1 (Lu et al., 2009). Firstly, image features are extracted to represent the image content. Then, both the image content and image features are encrypted and stored in the database. When images are retrieved, encrypted features can be compared directly without decryption. Finally, the images which have the similar encrypted features with the query image will be the outputs. Next, we will introduce the implementation details in the following parts.

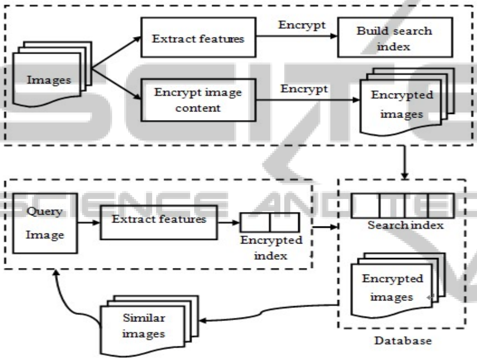

Figure 1: The framework of the image retrieval in the encrypted domain.

2.1 Image Feature Extraction

Image feature extraction is the key part of CBIR. In this paper, three kinds of low-level image features are extracted from images, including color feature, texture feature and shape feature.

2.1.1 Color Feature

In this paper, color histogram is extracted in HSV color space as the color feature.

Firstly, the RGB color space is converted into the HSV color space. Then, we quantize the color space into non-equidistant according to the color perception of human visual system. The hue H, the saturation S and the value V are divided into 8, 3 and 3 bins, noted as$\bar{H}$, $\bar{S}$ and$\bar{V}$, respectively. Then, these three components are integrated. The formula is

$$L = \bar{H}Q_SQ_V + \bar{S}Q_V + \bar{V} \tag{1}$$

where $Q_S = 3$ and $Q_V = 3$, the range of L is [0,71], representing 72 levels. Finally, we calculate the color histogram. Thus, we obtain a 72-dimensional color feature vector. (He, 2008)

2.1.2 Texture Feature

This paper applies the gray level co-occurrence matrix (Partio et al., 2002) to extract the texture feature.

Firstly, the gray level co-occurrence matrix, noted as $M_{(\Delta x,\Delta y)}(h,k)$, is obtained from the gray level image. Each element value of $M_{(\Delta x,\Delta y)}(h,k)$ is m_{hk}.We constitutes co-occurrence matrixes in four directions, $M_{(1,0)}$, $M_{(1,0)}$, $M_{(1,0)}$ and $M_{(1,0)}$. Then, four parameters (Angular Second Moment, Contrast, Entropy and Correlation) are calculated in the four co-occurrence matrixes. They are defined as follows:

$$ASM = \sum_h \sum_k (m_{hk})^2 \tag{2}$$

$$CON = \sum_h \sum_k (h-k)^2\, m_{hk} \tag{3}$$

$$ENT = \sum_h \sum_k m_{hk} log m_{hk} \tag{4}$$

$$COR = \frac{\sum_h \sum_k hkm_{hk} - \mu_x\mu_y}{\sigma_x\sigma_y} \tag{5}$$

where

$$\mu_\alpha = \sum_{x=1}^{h}\sum_{y=1}^{k} \alpha m_{xy}\,, \alpha = x,y \tag{6}$$

$$\sigma_\alpha{}^2 = \sum_{x=1}^{h}\sum_{y=1}^{k} m_{xy}(\alpha-\mu_\alpha)^2\,, \alpha = x,y \tag{7}$$

Finally, calculate the means and standard deviations of these four parameters by combining with the four directions: μ_{ASM},σ_{ASM}, μ_{CON}, σ_{CON}, μ_{ENT}, σ_{ENT}, μ_{COR} and σ_{COR} . Moreover, taking into account the difference of each part of image texture, we divide each image into four equal parts before extracting the gray level co-occurrence matrix. Thus, we obtain a 32-dimensional texture feature vector.

2.1.3 Shape Feature

In this paper, the Hu invariant moments (Hu, 1962) are extracted to represent the shape feature.

Suppose *f(i,j)* is the function of digital image. Its $(p+q)^{th}$ order moment is

$$M_{pq} = \sum_i \sum_j i^p j^q f(i,j), p,q = 0,1,2 \ldots \tag{8}$$

The central moment is

$$m_{pq} = \sum_i \sum_j (i-i')^p (j-j')^q f(i,j)\,, p,q = 0,1,2 \ldots \tag{9}$$

where $i' = M_{10}/M_{00}\,, j' = M_{01}/M_{00}$ .The $(p+q)^{th}$ normalized central moment is

$$\eta_{pq} = \frac{m_{pq}}{m_{00}^r}, r = \frac{p+q+2}{2} \tag{10}$$

Hu proposed seven moments as follow:

$$\begin{cases} \varphi_1 = \eta_{20} + \eta_{02} \\ \varphi_2 = (\eta_{20} - \eta_{02})^2 + 4\eta_{11}^2 \\ \varphi_3 = (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\ \varphi_4 = (\eta_{30} + \eta_{12})^2 + (\eta_{21} + \eta_{03})^2 \\ \varphi_5 = (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - \\ \qquad 3(\eta_{21} + \eta_{03})^2] + (3\eta_{21} - \eta_{03})(\eta_{21} + \\ \qquad \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\ \varphi_6 = (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] + \\ \qquad 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\ \varphi_7 = (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - \\ \qquad 3(\eta_{21} + \eta_{03})^2] - (\eta_{30} - 3\eta_{21})(\eta_{21} + \\ \qquad \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \end{cases} \tag{11}$$

These moments have invariance of transform, rotation and scaling. Thus, we can obtain a 7-dimensional shape feature vector.

In this paper, the gaussian normalization is applied before similarity comparison. After the gaussian normalization, the values of feature vector are ranged within $[-1,1]$. Therefore, we can effectively eliminate the difference which generates due to the different range of the values in similarity comparison, and make the weight coefficients of the different components of the same feature roughly the same.

Finally, integrating the three kinds of feature vectors together, we can get a 111-dimensional feature vector to represent the image content.

2.2 Image Feature Encryption

In this paper, AES and other encryption algorithms are used to encrypt the image content, and will not be described here due to the limited paper space. Next, we will describe the image feature encryption method, which is the key component in our proposed secure image retrieval scheme.

The CBIR is based on the distance between the image features. In order to achieve the goal of comparing the encrypted image features without decryption, this paper applies an algorithm called bit-plane randomization (Lu et al., 2009) which can preserve the distance between the encrypted image features. In this paper, we preserve the highest bit-planes and encrypt the subsequent two bit-planes. The so-called bit-plane is the bits which have the same weight. Clearly, different bit-planes have different significance. The most significant bits (MSB) have the most importance of value.

The specific process of encryption is as follow: First of all, the values of the feature vector have been normalized into [-1,1] after gaussian normalization. In order to extract bit-planes, we use the following equation to process all the values:

$$e' = (e + 1) * 100 \tag{12}$$

where *e'* and *e* represents the processed and original values, respectively. Then, the range of all the values of the feature vector is [0,200]. Due to $200_{(10)} = 11001000_{(2)}$, we can get 8 bit-planes. Next, a random binary string is generated as the cryptographic key. After that, the first five MSBs are extracted. The fourth and the third bit-planes are XORed with the cryptographic key respectively, while the highest three bit-planes (7-5) are preserved. Finally, the binary values should be transformed back into decimal notation. Thus, the values of the image feature vector can be protected efficiently.

2.3 Similarity Comparison

In this paper, both L1 and L2 distance are used to perform similarity comparison. L1 and L2 distance between *N*-dimensional vectors can be defined as

$$D_1 = \sum_{i=1}^{N} |A_i - B_i| \tag{13}$$

and $$D_2 = \sqrt{\sum_{i=1}^{N} (A_i - B_i)^2} \tag{14}$$

3 EXPERIMENTS AND ANALYSIS

To validate the effectiveness of the proposed secure image retrieval, we conduct an experiment on an image database which contains two image datasets:

- The Corel image database: this database contains 1000 color images which are classified into 10 categories (100 images in each category) images, i.e. African, Beach, Architecture, Buses, Dinosaurs, Elephants, Flowers, Horses, Mountain and Food. Image sizes are either 256*384 or 384*256. This database has been widely used in evaluating color image retrieval;
- The existing public datasets: we select 1232 images from the existing public datasets, such as MSRC and Flickr. In this database, images are classified into 8 categories, i.e. Bicycles, Buses, Cars, Cats, Cows, Motorbikes, People and Sheep. Image sizes are not fixed. Most of the images are social images and are generally used for image segmentation or recognition.

In this database, the total number of images is 2232. After merging the similar classifications, we get 16 categories, i.e. Bicycles, Cars, Cats, Cows, Motorbikes, People, Sheep, Beach, Architecture, Buses, Dinosaurs, Elephants, Flowers, Horses, Mountain and Food.

In this experiment, we evaluate the performances of the cipertext retrieval and the plaintext retrieval by precision-recall curve. Moreover, the performances of retrieval based on L1 distance and L2 distance are also evaluated. The precision-recall curves are shown in Figure 2, and the results of plaintext and ciphertext retrieval are shown in Figure 3 and Figure 4, respectively.

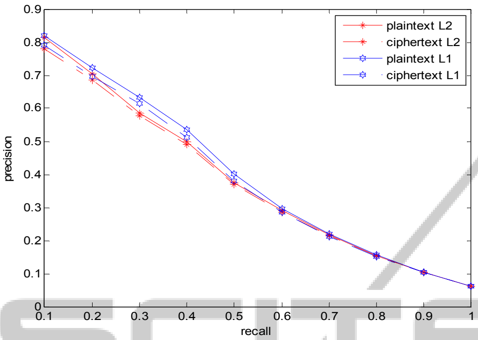

Figure 2: Retrieval performance: ciphertext and plaintext retrieval on the mixed image database.

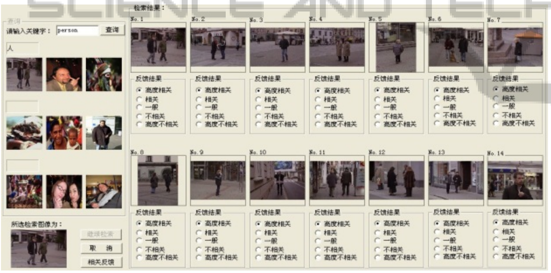

Figure 3: The results of plaintext retrieval.

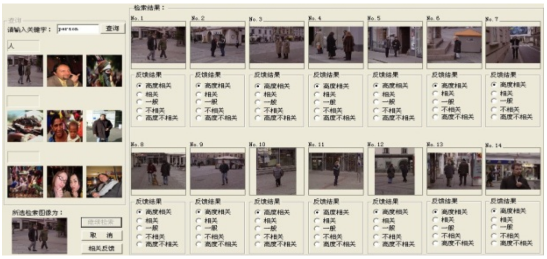

Figure 4: The results of ciphertext retrieval.

It can be seen from Figure 2 that, the proposed secure image retrieval scheme can achieve the comparable retrieval performance with plaintext retrieval. When we compare the performance of retrieval based on L1 distance with that based on L2 distance, both of the methods can achieve a good performance in the encrypted domain while L1 distance performs slightly better than L2. Obviously, the scheme proposed in this paper can achieve good performance while preserving the image formation security at the same time.

4 CONCLUSIONS

This paper proposes a secure image retrieval scheme in the encrypted domain. When images are retrieved, encrypted features can be compared directly without decryption. The experimental results show that this secure image retrieval scheme could achieve the comparable retrieval performance, while preserving the image formation security at the same time. Future work will extract better features, like SIFT and other local invariant features, to further improve the retrieval performance.

ACKNOWLEDGEMENTS

The work in this paper is supported by Program for New Century Excellent Talents in University (No.NCET-11-0892), Doctoral Fund of the Ministry of Education, the National Natural Science Foundation of China (No.61003289,No.61100212) , the Natural Science Foundation of Beijing (No. 4102008), the Excellent Science Program for the Returned Overseas Chinese Scholars of Ministry of Human Resources and Social Security of China, Scientific Research Foundation for the Returned Overseas Chinese Scholars of MOE, Youth Top-notch Talent Training Program of Beijing Municipal University, the Fundamental Research Funds for the Central Universities (No.NS2012045).

REFERENCES

Meng Wang, Kuiyuan Yang, Xian-Sheng Hua, Hong-Jiang Zhang, 2010. Towards a Relevant and Diverse Search of Social Images. IEEE Transactions on Multimedia, vol. 12, no. 8, pp. 829-842.

Gao, Y., Wang, M., Zha, Z. J., Shen, J. L., Li., X. L., Wu, X. D, 2012. Visual-Textual Joint Relevance Learning for Tag-Based Social Image Search. IEEE Transactions on Image Processing. In press.

M. Zeghid, M. Machhout, L. Khriji, A. Baganne, and R. Tourki, 2007. A Modified AES Based Algorithm for Image Encryption. International Journal of Computer Science and Engineering Volume1 Number1,pp. 70-75.

Z. Erkin, A. Piva, S. Katzenbeisser, R. L. Lagendijk, J. Shokrollahi, G. Neven, and M. Barni, 2007. Protection and retrieval of encrypted multimedia content: when cryptography meets signal processing. EURASIP Journal on Information Security, vol. 7, no. 2, pp. 1-20.

J. Eakins, M. Graham, 1999. Content-based image retrieval. Tech. Rep. JTAP-039, JISC Technology Application Program. Newcastle upon Tyne.

W. Lu, A. Swaminathan, A. L. Varna, and M. Wu, 2009.

Enabling search over encrypted multimedia databases. Proc. of SPIE, vol. Media Forensics and Security.

W. Lu, Varna, A. L., Swaminathan, A., and Wu, M., 2009. Secure image retrieval through feature protection. IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 1533–1536.

Lin He, 2008. A Research of The User's Interest Used in Image retrieval. Signal and Information Processing Lab, Beijing University of Technology.

M. Partio, B. Cramariuc, M. Gabbouj, and A. Visa, 2002. Rock texture retrieval using gray level co-occurrence matrix. In 5th Nordic signal processing symposium on board Hurtigruten. Norway.

Ming-Kuei Hu, 1962. Visual pattern recognition by moment invariants. *IEEE Transactions on Information Theory*, vol. 8, no. 2, pp. 179–187.