Secure Content-Based Image Retrieval in Cloud Computing

Swati B Daberao, V.S. Mahalle
Department of Computer Science and Engineering
Shri Sant Gajanan Maharaj College of Engineering Shegaon, India.

***Abstract*— Cloud based storing visual data have increased in recent years, following the emergence of many interactive multimedia services and applications for mobile devices in personal and business scenarios. This was a determining factor in the adoption of cloud-based data outsourcing solutions. However, even outsourcing cloud data storage brings new security challenges that need to be addressed carefully. We offer a secure framework for storing and recovering outsourced privacy protection in large shared image files. Our proposal is based on CBIR, a new image cryptography scheme that features image recovery based on content. The framework allows both encrypted storage and querying for content-based image retrieval, while maintaining privacy in the face of honest but curious cloud administrators. We have built a prototype of the proposed framework, analyze and formally test its safety properties, and experimentally evaluate its performance and recovery accuracy. Our results show that CBIR is probably safe, enabling more efficient operations than existing proposals, both in terms of time and space complexity, and opens the way for new scenarios of practical application.**

Keywords –.Cloud computing, Data and computation outsourcing, encrypted data processing, searchable encryption content based image retrieval.

I. INTRODUCTION

Content-based image retrieval (CBIR) is also known as query by image content and content-based visual information retrieval is the use of computer vision to the image retrieval problem of searching for digital images in large databases. "Content-based" means that the search will analyze the actual contents of the image. The term 'content' in this context might refer colors, shapes, textures, or any other information that can be derived from the image itself. Without the ability to examine image content, searches must rely on metadata such as captions or keywords. Such metadata must be generated by a human and stored exactly each image in the database. An image retrieval system returns a set of images from a collection of images in the database to meet users demand with similarity assessment such as image content similarity, edge pattern similarity, color similarity etc. Image retrieval system offers an efficient way to access, browse, and retrieve a set of similar images in the real-time applications. As a result of recent advancements in digital storage technology, it is now possible to create large and extensive databases of digital imagery. These collections may contain millions of images and terabytes of data. For users to make the most of these databases effective, efficient methods of searching must be devised. Prior to automated indexing methods, image databases were indexed according to keywords that were both decided upon and entered by a human categorizer. Unfortunately, this practice comes with two very severe shortcomings. First, as a database becomes increasingly large the manpower required to index each image becomes less practical. Secondly, two different people, or even the same person on two different days, may index similar images inconsistently. The result of these inefficiencies is a less than optimal search result for the end user of the system.

Computer do the indexing based on a CBIR scheme attempts to address the shortcomings of human-based indexing. Since a computer can process images at a much higher rate, while never tiring For example, each CBIR system needs to be tuned for its particular use in order to give optimal results. A retrieval system designed for querying medical x-ray images will more than likely prove to be a poor system for retrieving satellite images of South American rain forests. In addition, presently employed algorithms cannot yet consistently extract abstract features of images, such as emotional response, that would be relatively easy for a human to observe. Several approaches have been developed to capture the information of image contents by directly computing the image features from an image. The image features are directly constructed from the typical Block Truncation Coding or half toning based compressed data stream without performing the decoding procedure. These image retrieval schemes involve two phases, indexing and searching, to retrieve a set of similar images from the database. The indexing phase extracts the image features from all of the images in the database which is later stored in database as feature vector. In the searching phase, the retrieval system derives the image features from an image submitted by a user.

II. Related Work

Y. Gong and S. Lazebnik proposed the problem of learning binary codes that preserves the similarity for an efficient search for similarity in large-scale image collections is formulated by terms of zero-rotation data centering to minimizing quantization error by mapping data to the vertices of a zero-center binary hypercube as well as proposing a simple and efficient alternative minimizing algorithm to perform this operation [1].

The author Y. Pan, T. Yao, T. Mei, H. Li, C.-W. Ngo, and Y. Rui**,** proposed an approach for jointly exploring cross-view learning and the use of click data. The cross view learning is used for creating latent subspace with the ability to compare information from incomparable original views (ie text and image views), and use of click data explores access data that is widely available and freely accessible for understanding of the query [2].

The author D. Zhai, H. Chang, Y. Zhen, X. Liu, X. Chen, and W. Gao have been proposed HFL for the searching of inter-vision similarities. A new multimode HFL method, called Parametric Local Multimodal Hashing (PLMH) that can learn a set of hash functions to adapt locally to the data structure of each mode [3].

author G. Ding, Y. Guo, and J. Zhou proposed the problem of learning hash functions in the context of multimodal data for the search for similarity between cross-views is formulated by they proposed the Collective Matrix Factorization Hashing (CMFH) method which can generates unique hash codes for various modalities of single instance through collective matrix factorization along with the latent factor model [4].

Author H. Jegou, F. Perronnin, M. Douze overcommed the problem of large-scale image search. For this purpose they have provided three restrictions i.e search accuracy, effciency and memory usage and proposed different ways to add local image descriptors into a vector and demonstrated that Fisher's kernel performs as much better as visual bag approach for any given vector dimension [5].

The author J. Zhou, G. Ding, and Y. Guo proposed a new LSSH (Latent Semantic Sparse Hashing) algorithm to perform a search for similarity between modes using Sparse Coding and Matrix Factorization. For this purpose LSSH uses Sparse Coding to acquire the most important image structures and Matrix Factorization to learn the latent concepts of the text. [6].

The author Z. Yu, F. Wu, Y. Yang, Q. Tian, J. Luo, and Y.Zhuan proposed a Discriminative Coupled Dictionary Hashing (DCDH), in which the paired dictionary for each mode is acquired with secondary information (for example, categories). These coupled dictionaries not only preserve the intra-similarity and interconnection between multimode data, but also contain dictionary atoms that are semantically discriminating (that is, data in the same category are reconstructed from atoms in the similar dictionary) [7].

The author H. Zhang, J. Yuan, X. Gao, and Z. Chen has been proposed a method of cross-media recovery based on short and long-term relevance feedback. This method focused on two typical types of multimedia data, i.e. image and audio. Firstly they have created a multimodal representation through a statistical correlation between the image arrays and audio entities, and they defined the metric of the distance between the means for the measurement of similarity; therefore an optimization strategy based on relevant feedback combines the results of short-term learning and long-term accumulated knowledge in the objective function [8].

The author A. Karpathy and L. Fei-Fei proposed a model generating the descriptions of natural language of images and their regions. This approach have advantage of image data sets and their sentence descriptions to know the intermodal correspondences between language and visual data.The alignment model is based on combination of convolutional neural networks on image regions, bidirectional recurrent neural networks on sentences . The structured goal aligns two modalities through a multimodal model [9].

The author J. Song, Y. Yang, Y. Yang, Z. Huang, and H. T. Shen proposed a multimedia recovery paradigm to innovate large-scale research of different multimedia data. It is able to find results from different types of media of heterogeneous data sources, for example by using a query image to retrieve relevant text documents or images from different data sources [10].

III. Proposed Approaches:-

We propose a secure framework for the storage and recovery of the subcontracted privacy protection in large archives of shared images. Our proposal is based on CBIR, a novel Encryption scheme of the image that presents image recovery properties based on content. The framework allows both encrypted storage and search using content-based image retrieval queries while preserving privacy against honest but curious cloud administrators. We have built a prototype of the proposed framework, formally analyzed and tested its safety properties, and experimentally assessed its performance and accuracy of recovery. Our results show that CBIR is probably safe, allowing more efficient operations that the existing proposals, both in terms of complexity of time and space, and opens the way to new scenarios of practical application.

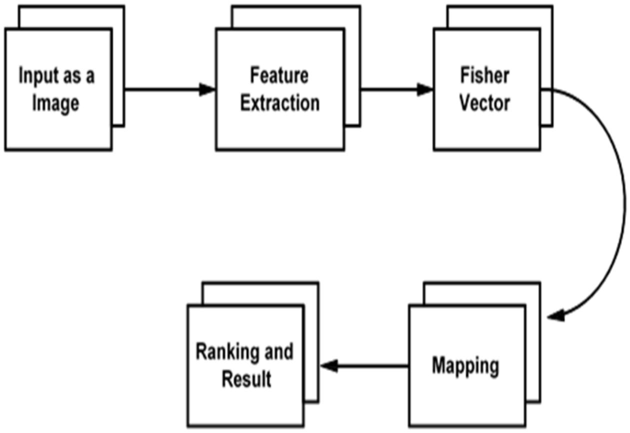


Fig.

Flow Diagram

Algorithms:

1. **AES Encryption Algorithm**

AES (advanced encryption standard).It is symmetric algorithm. It used to convert plain text into cipher text .The need for coming with this algorithm is weakness in DES. The 56 bit key of des is no longer safe against attacks based on exhaustive key searches and 64-bit block also consider as weak. AES was to be used128-bit block with128-bit keys.
Rijendeal was founder. In this drop we are using it to encrypt the data owner file.
Input:
128_bit /192 bit/256 bit input (0, 1)
Secret key (128_bit) +plain text (128_bit).
Process:
10/12/14-rounds for-128_bit /192 bit/256 bit input
X or state block (i/p)
Final round: 10, 12, 14
Each round consists: sub byte, shift byte, mix columns, add round key.
Output:
Cipher text (128 bit)

2. **Global Color Algorithms**

The Global color is used to represent Images and it also extract the key points from the number of images. The Global color then calculates the descriptors of the extracted key points and a set of variable-sized key points in the Global color space represents a particular image.

Steps:

1. The procedure to search in a repository R with query image Q.
2. The input for this operation on the user side is IDR, Q, repository key rkR, and parameter k (the number of most similar results to be returned).
3. User U starts by generating Q's searching trapdoor CQ, through IES-CBIR.
4. Then sends it to the cloud server, along with k and IDR, as parameters for the Search remote invocation.
5. The cloud starts by extracting CQ's feature-vector, stems it against CBR to determine its visual words vwCQ, and accesses IdR with them to retrieve the respective posting lists PLvw.
6. Then, for each image referenced in each of the posting lists retrieved, the cloud calculates its scaled tf-idf score and adds it to the set of results for the query. In this set, scores for the same image but different visual word are summed.
7. Finally, the cloud sorts this set by descending score and returns the results to user.

Dataset:

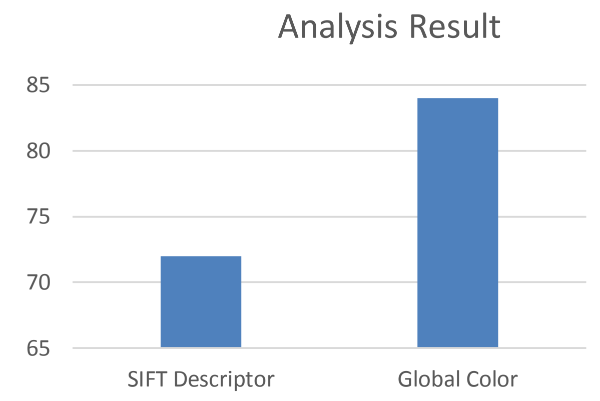
This is a 10 class land use image dataset meant for research purposes.
There are 100 images for each of the following classes:

- Aboriginal
- Beach
- Historical Places
- Buses
- Kangaroo animal
- Elephant animal
- Roses
- Horse
- Mountain
- Recipe Dish

Each image measures 256x256 pixels.
The images were extracted from large images from the Flicker professional collection for various areas around the country. The pixel resolution of this public domain imagery is 1 foot.

Result Analysis:

Comparison Graph



Comparison table

	IES-CBIR	SSE
Accuracy Percentage	84%	72%

The experimental result evaluation, we have notation as follows:
TP: True positive (correctly predicted number of instance)
FP: False positive (incorrectly predicted number of instance),
TN: True negative (correctly predicted the number of instances as not required)
FN false negative (incorrectly predicted the number of instances as not required),
On the basis of this parameter, we can calculate four measurements
Accuracy = TP+TN÷TP+FP+TN+FN
Precision = TP ÷TP+FP
Recall= TP÷TP+FN
F1-Measure = 2×Precision×Recall ÷Precision+ Recall

Conclusion

In this Paper, we have proposed a new secure framework for the external storage of privacy protection, research and recovery of large-scale dynamic image archives, where the reduction of the general expenses of the customer is central appearance. At the base of our framework there is a new cryptography scheme, specifically designed for images, called CBIR. The key to its design is the observation that in the images, color information can be separated from the plot information, allowing the use of different cryptographic techniques with different properties for each and allowing to preserve privacy Image recovery based on the content that will be created from unreliable third-party cloud servers. We formally analyze the safety of our proposals and further experiments the evaluation of the implemented prototypes revealed that our approach reaches an interesting exchange between precision and I remember in the CBIR, while exhibiting high performances and scalability compared to alternative solutions.

REFERENCES

[1] Y. Gong, S. Lazebnik, A. Gordo, and F. Perronnin, “Iterative quantization: A procrustean approach to learning binary codes for large-scale image retrieval,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 35, no. 12, pp. 2916–2929, Dec. 2013.

[2] Y. Pan, T. Yao, T. Mei, H. Li, C.-W. Ngo, and Y. Rui, “Clickthrough-based cross-view learning for image search,” in Proc. 37th Int.ACMSIGIR Conf. Res. Develop. Inf. Retrieval, 2014, pp. 717–726.

[3] D. Zhai, H. Chang, Y. Zhen, X. Liu, X. Chen, and W. Gao, “Parametric local multimodal hashing for cross-view similarity search,” in Proc. 23rd Int. Joint Conf. Artif. Intell., 2013, pp. 2754–2760.

[4] G. Ding, Y. Guo, and J. Zhou, “Collective matrix factorization hashing for multimodal data,” in Proc. IEEE Conf. Comput. Vis. Pattern Recog., 2014, pp. 2083–2090.

[5] H. J_egou, F. Perronnin, M. Douze, J. S_anchez, P. P_erez, and C. Schmid, “Aggregating local image descriptors into compact codes,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 34, no. 9, pp. 1704–1716, Sep. 2011.

[6] J. Zhou, G. Ding, and Y. Guo, “Latent semantic sparse hashing for cross-modal similarity search,” in Proc. 37th Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval, 2014, pp. 415–424.

[7] Z. Yu, F. Wu, Y. Yang, Q. Tian, J. Luo, and Y. Zhuang, “Discriminative coupled dictionary hashing for fast cross-media retrieval,” in Proc. 37th Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval, 2014, pp. 395–404.

[8] H. Zhang, J. Yuan, X. Gao, and Z. Chen, “Boosting cross-media retrieval via visual-auditory feature analysis and relevance feedback,” in Proc. ACM Int. Conf. Multimedia, 2014, pp. 953–956.

[9] A. Karpathy and L. Fei-Fei, “Deep visual-semantic alignments for generating image descriptions,” in Proc. IEEE Conf. Comput. Vis. Pattern Recog., Boston, MA, USA, Jun. 2015, pp. 3128–3137.

[10] J. Song, Y. Yang, Y. Yang, Z. Huang, and H. T. Shen, “Inter-media hashing for large-scale retrieval from heterogeneous data sources,” in Proc. Int. Conf. Manage. Data, 2013, pp. 785–796.