International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 04 | Apr 2020 www.irjet.net p-ISSN: 2395-0072

A Content Based Image Retrieval Scheme Using Bag-Of-Encrypted Words

P.Mohamad Suhail[1], L.Karthik Raja[2], Ch.Charan[3] ,J.K.Arun Nehru[4]

[1,2,3]B.Tech Student, Computer Science Engineering, SRM IST,Chennai

[4]Assitant Professor, Computer Science Engineering, SRM IST,Chennai

---

Abstract - *With the exponential growth of digital images, the content-based Image Retrieval techniques have been extensively studied. CBIR service is typically very costly in terms of computing and storage resources. Therefore, outsourcing the CBIR service to the cloud server, which is equipped with massive resources, is a smart choice. Data security is a major problem, however, because the cloud service cannot be totally positive. In this we propose a CBIR outsourced scheme based on a concept of a novel bag- of-encrypted-words (BOEW). The picture is encrypted by substitution of the color value, block permutation and permutation of the intra-block pixels. Then, the cloud server computes the local histograms from the encrypted image fragments. All the local histograms are grouped together, and the centers of the clusters are used as the encrypted visual words. The bag-of-encrypted-words (BOEW) model is thus constructed to represent each image by a function vector, i.e., a structured histogram of the encrypted visual terms. The resemblance between images can be determined directly by the distance between feature vectors on the side of the cloud server in Manhattan. Experimental results and safety review on the proposed scheme show its accuracy and protection of the search.*

***Key Words*: CBIR , Encrypted Image,BOEW,Cloud,Histogram**

1. INTRODUCTION

A BOEW model for outsourcing CBIR is proposed. We suggest encrypting blocks of images and ensuring that the stable and useful local features can be extracted directly from the encrypted blocks. To generate the encrypted visual terms, the k-means clustering algorithm is used. The final function vectors are then constructed with the visual terms, including the encrypted ones. The final function vectors are then constructed with the visual terms, including the encrypted ones. Euclidean or Manhattan distance can directly calculate the similarity between the feature vectors. The growing cloud industry has provided a service paradigm that helps to reduce the burden on users of maintaining IT infrastructure and reduce costs for both businesses and individual users. Cloud computing is a modern type of Internet-based computing that provides on demand computers and other devices with shared compute processing resources and data. This is the delivery of resources accessed over the internet. Cloud computing services can be public, hybrid or private. In addition to the immense advantages of CBIR outsourcing, image protection is now the image owner's greatest concern. Both the image and query image database should be properly secured.

2. RELATED WORK

CBIR techniques have been researched for over 20 years and demonstrated its effectiveness in many real-world applications. Lu et al suggested the first CBIR scheme that would maintain privacy over the encrypted image database. The scheme used the collection of visual expressions to represent images. Jaccard's distance between the sets of visual words calculated the similarity between the images. However, due to privacy issues it cannot be outsourced directly to the cloud. It is noteworthy that the image features will also leak information about image contents in addition to the original image data, if they are not well secured. To secure

the visual words the min-hash algorithm and order-preserving encryption are used. Three image feature security strategies including bitplane randomization, random projection, and randomized unary encoding were explored in another study by Lu et al. The bitplane randomisation and random unary encoding help Hamming distance calculation in the encryption domain. The random projection in the encryption domain supports the estimated measurement of the distance L1. Lu et al compared the three methods listed with the homomorphic encryption and indicated that the homomorphic encryption consumed much more Yuan et al computational and communications resources protected the image features with local sensitive hashing and Cuckoo Hashing to help secure search for similarity. This method has been used to discover the social connections among image owners. Xia et al suggested a CBIR privacy-conserving scheme based on Scale-Invariant Feature Transform (SIFT) and Earth Mover's Distance (EMD) functions. The EMD calculation is simply a linear problem with the program. The linear transformation was used during the EMD problem solving process to protect the privacy information.Yuan et al developed an encrypted image search scheme based on the secure algorithmkNN (k- nearest neighbors) and created a tree index to boost the search performance. Chen et al have proposed a retrieval scheme based on the Markov method over encrypted files. The integrity of the images was secured by encrypting the Huffman table into JPEG files. The Markov characteristics were extracted directly from the DCT coefficients that were decoded with the encrypted Huffman table. In this scheme, random permutation preserves the color values, and the pixels are shuffled by rows and columns. Upon uploading the encrypted image to the cloud server the color histograms of the HSV (Hue-Saturation-Value) are extracted from the encrypted images on the cloud server side. The differences between images can be calculated directly by the Hamming distances of the respective histograms. In this way, the owner of the image only undertakes the encryption of the file. Certain functions are outsourced to the cloud server, such as extraction of functionality, index generation and search process. However, the global histogram is too stiff for problems with image retrieval. Liu et al have sought to improve the precision of the recovery with a histogram of variance. Yet the change is not significant. In this paper we are proposing an outsourced CBIR scheme focused on stable local characteristics.

The suggested scheme also outsources the tasks of extracting functions, building indexes and searching to the cloud server, while at the same time achieving much greater accuracy of recovery by using the proposed BOEW model.

3. EXISTING SYSTEM

In recent years, storage demands for visual data have increased, following the advent of various highly interactive multimedia platforms and mobile device applications in both personal and corporate scenarios. Existing ideas in this area remain largely unworkable, including those involving complete homomorphism encryption, which is still too costly in computational terms. Considering that mobile customers usually have limited computing and storage capacity, they continue to rely on cloud providers to store and process bulky data such as images. In this case, mobile clients (users) want to assign storage of their private image archives to a cloud provider, while dealing with the limitations of storage space, computing power, and battery life of their devices.

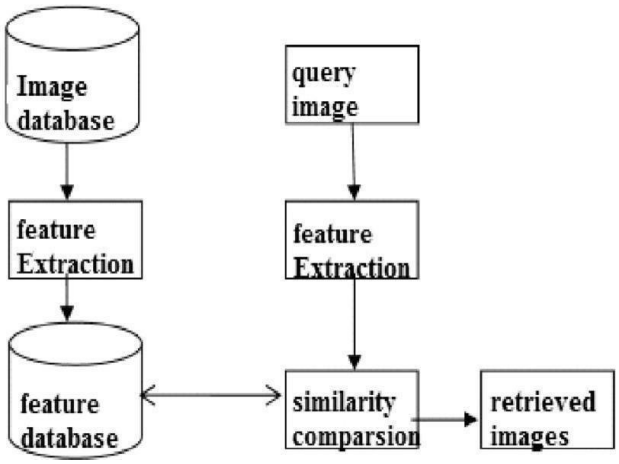

Fig 1. Existing System

4. PROPOSED SYSTEM

Our proposal is based on IES-CBIR, a new Image Encryption Scheme that displays information-related Image Retrieval properties. The system allows both encrypted storage and search queries using content-based image retrieval. Images are outsourced to cloud-inhabited servers. Multiple users use each repository to add their own photos and/or to search using a query image. Every server is a single user. Once a repository is built, the user creates a new repository key and then shares it with other trusted users, allowing them to search in the repository and add / update photos. For this work we use the representation Bag-Of- Encrypted-Words (BOEW) to construct a vocabulary tree for each repository and an inverted list index. We choose this approach for indexing because it demonstrates strong search efficiency and properties of scalability. Throughout the BOEW model, feature-vectors are hierarchically clustered into a vocabulary tree (also known as codebook), where each node denotes a representative feature-vector in the array, and the most representative nodes (called visual words) are selected from the leaf nodes.

5. SYSTEM ARCHITECTURE

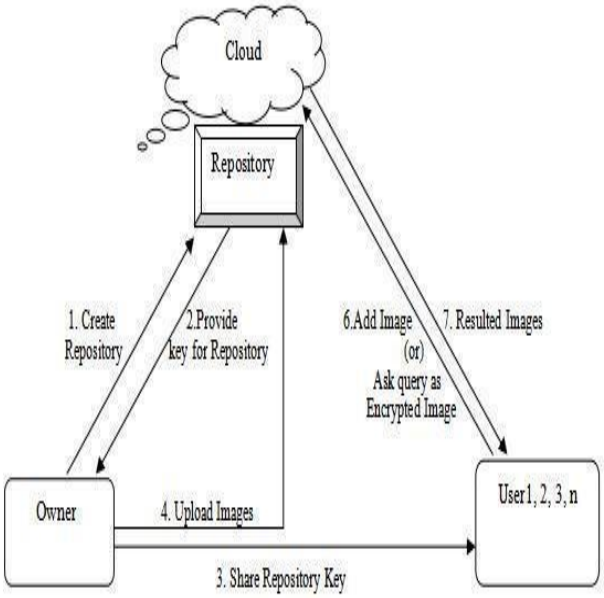

Fig 2. System Architecture

6. MODULES

6.1 Create Repository & Upload Images

Repository is data-gathering storage space. Each repository is created by individual user. He is the proprietor of the repository. Then, by using the RSA algorithm, he creates a key for that server, and shares it with users who all have an account to access it. Now, multiple users can access the Repository with an owner's permission. Instead, the user uploads huge quantities of image datasets into the cloud as a zip file.

6.2 Codebook & Index generation

The cloud administrator has the task of producing image-based documents that are useful for users to scan images. So, he extracts zip file and the technique of CBIR encryption. This encrypts images based on color values and texture characteristics, and even shuffles the pixels in both column-wise and row-wise. Instead, for those encrypted files, he creates codebook, index and image key. These files are used to improve cloud search efficiency and also manage the time properly when retrievinganswers.

6.3 Add Image/Query to cloud:

Users can also use the cloud and get their own photos added to the server. Therefore, if there is a ' n ' number of users in that cloud, then the server has a chance to expand quickly. Currently, the archive has accumulated ' n ' number of images in different domains. For security purposes all files are stored in encrypted format. Instead, in cloud, the user will ask for question. The user has to ask question in the encrypted image format for using CBIR encryption technique.

6.4 Content Based Searching & Retrieval:

The cloud extracts the features of an original file, after obtaining encrypted file query. Now add content-based search to the codebook and image index by using the extracted functions. Obviously, an authenticated image would be the search results now. This resulting response will be sent to the corresponding user. Now, the user may use CBIR decryption technique to decrypt the photos that have been retrieved. So, because of the large dataset the answer will be very good and delicious.

7. DISCUSSION

In many ways the proposed scheme can be strengthened. Firstly, better local features can be extracted under the proposed BOEW model. This depends on how they process and encrypt the intra-pixels. For example, as local features one can extract and encrypt the gradient information from image blocks. Using texture information may achieve better accuracy in retrieval. The leakage of information is an issue to overcome, however. Secondly, our scheme is based on uncompressed pictures. Practical problem might be storage consumption. According to the encrypted image may also be compressed to reduce the cost of storage given loss of the connection between the pixels. In this, we don't mention the compression of images much. Finally, applying the BOEW model to retrieval of encrypted JPEG images may be a meaningful task. The proposed scheme is parameters- strong. Furthermore, even though a very small part of local histograms are used for clustering, our scheme can get satisfying retrieval accuracy.

8. EXPERIMENTAL RESULTS

Repository is data-gathering storage space. Each repository is created by individual user. He is the proprietor of the archive. Instead, by using the RSA algorithm, he creates a key for that server, and shares it with users who all have an account to access it. Now, multiple users can access the Repository with an owner's permission. Instead, the user uploads large amounts of image datasets into the cloud as a zip file. The cloud administrator has the task of producing image-based documents that are useful for users to scan images. So, he extracts zip file and the technique of CBIR encryption. It encrypts images based on color values and texture characteristics, and also shuffles the pixels in both column-wise and row-wise. Then, for those encrypted files, he creates codebook, index and image key. These files are used to improve cloud search efficiency and also manage the time properly when retrieving answers. Users will also use the cloud and get their own images added to the server. And, if there is a' n' number of users in that cloud, then the repository has a chance to expand quickly. Currently, in various contexts, the repository has set of' n' number of images. All files are stored for security purposes in encrypted format. Then, in cloud, the user will ask for question. The take question is in encrypted image format using CBIR encryption technique. The cloud extracts the features of an original file, after obtaining encrypted file query. Now add content-based search to the codebook and image index by using the extracted functions. Obviously, an

authenticated image would be the search results now. The resulting response will be submitted to the corresponding person. Now, the user can use CBIR decryption technique to decrypt the photos that have been retrieved. So, because of the large dataset the final answer will be seen.

9. SCREENSHOTS

Fig 3: Website

Fig 4; Registration Page

10 . CONCLUSION

A new, privacy-conserving CBIR scheme is proposed in this paper. A novel Bag-of-Crypted-Words (BOEW) model is designed to achieve good accuracy in the retrieval. As a case study, color value substitution, block permutation, and intra-block pixel permutation protect the image material. Regional histograms are measured as being regional characteristics. We propose a modern, privacy-conserving CBIR scheme. A novel Bag-of-Crypted-Words (BOEW) model is designed to provide good accuracy in the recovery process. As a case study, the image content is covered by color value substitution, block permutation, and intrablock pixel permutation. The regional histograms are measured as regional features. Thus, a stable and secure system for outsourced privacy-conserving storage and retrieval has been successfully introduced in broad shared image repositories.

11. REFERENCES

[1] C. S. Lu, "Homomorphic encryption-based secure sift for privacy- preserving feature extraction," *Proceedings of SPIE The International Society for Optical Engineering*, vol. 7880, no. 2, pp. 788 005–17, 2011.

[2] B. Ferreira, J. Rodrigues, J. Leit˜ao, and H. Domingos, "Privacy- preserving content-based image retrieval in the cloud," in *IEEE 34th Symposium on Reliable Distributed Systems*. IEEE, 2015, pp. 11–20.

[3] B. Ferreira, J. Rodrigues, J. Leitao, and H. Domingos, "Practical privacy-preserving content-based retrieval in cloud image repositories," *IEEE Transactions on Cloud Computing*, vol. PP, no. 99, pp.1–1, 2017.

[4] Y. Rui, T. S. Huang, M. Ortega, and S. Mehrotra, "Relevance feed- back: a power tool for interactive content- based image retrieval," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 8, no. 5, pp. 644–655, 1998.

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

[6] C. B. Akg¨ ul, D. L. Rubin, S. Napel, C. F. Beaulieu, H. Greenspan, and B. Acar, "Content-based image retrieval in radiology: current status and future directions," *Journal of Digital Imaging*, vol. 24, no. 2, pp. 208–222, 2011.