International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 06 Issue: 05 | May 2019 www.irjet.net p-ISSN: 2395-0072

IMPLEMENTATION OF PRIVACY PRESERVING CONTENT BASED IMAGE RETRIEVAL IN CLOUD

Athira Nair M[1], Asha Vijayan[2]

[1]Post graduation student, Dept of CS, College of Engineering Kidangoor, Kerala, India
[2]Assistant professor, Dept of CS, College of Engineering Kidangoor, Kerala, India

---

Abstract - *The importance and availability of large amount of images enforced the development of Content based image retrieval (CBIR) or Query based image retrieval (QBIR) systems. CBIR retrieves all the images similar to a given query image. As CBIR demands high storage space, images are outsourced to the cloud server. In order to preserve privacy of images in cloud, images are encrypted and watermarked before outsourcing. In this paper, for feature extraction FAST or feature from accelerated segment test algorithm is used. Advanced encryption scheme (AES) is used for encrypting images and least significant bit (LSB) is used for watermarking images before outsourcing images to cloud. Similar images are obtained using k nearest neighbour (knn) algorithm.*

Key Words: **AES, CBIR, FAST, LSB**

1. INTRODUCTION

The amount of images in our daily life is increasing widely everyday. In medical field millions of digital images especially in radiology department is being generated. Many social networking sites such as hello, facebook, instagram etc. uploads millions of images. NASA's earth observing system generates terabytes of images [1] [2]. This enormous generation of images forced the study of content based image retrieval systems (CBIR). In CBIR image owner holds database of images and whenever user provides a query image, system provides all the images similar to the given query image. But CBIR demands large storage space as images consumes more space compared to the text documents [3]. Additionally a large database may contain millions of images and also a single image may have high dimensionality. So it is better to store images in cloud server. Despite of the storage benefits of the outsourcing there are privacy issues. Cloud server is considered to be semi honest, that is honest but curious and tries to retrieve sensitive information. Many malicious administrator working for cloud having full access to the data stored in cloud.

To protect the images from such malicious administrator and external hackers all the images are encrypted using advanced encryption (AES) before storing in cloud. Images are also watermarked using least significant bit (LSB) technique for copyright protection. Raw image can't be used directly in my image processing applications, as most of the information in the system is redundant. Only expressive representation of the most relevant information is extracted. The process of finding expressive representation is called feature extraction. In this paper colour and shape feature are extracted from images. All the images in cloud are trained that is their shape or corner points are extracted using feature from accelerated test or FAST algorithm. Whenever an image owner uploads the image in addition to corner points it's colour feature is extracted. For colour feature extraction colour structure ratio, RGB ratio and edge histogram ratio is used.

When image user provides a query image, it's corner points are generated using FAST whose output is a text file containing large number of points . A limit of (+5,-5) is applied to each of this points. Each training image shape points are checked to see whether it is the neighbor of or included in the corner point of query image. All such images are obtained and it's average matching ratio and type is calculated. Images having highest matching ratio will be the result of the query image. Only the image name and it's owner's name will be visible to the image user if there is a similar image for a given query image. Image user has to send a request to the image owner to obtain the image. Image owner can either accept or reject the request. A user can be both an image owner or image user. If a user tries to upload the same image obtained from other user then it won't be possible because the image is watermarked by it's owner using LSB. In LSB, least significant bits of each pixel is replaced with bits of the secret data.

2. SYSTEM ARCHITECTURE

Admin add users to the system. A user can act as both image owner and image user. Image owner extract features from images. Image owner encrypt all the images and watermark them before outsourcing to the cloud server. When an image user provides a query image, system find similar images. If similar images are available then it provides the image name, description about the image and it's owner's name to the image user [4]. When image user sends request to the image owner for the image owner can either accept or reject the request. If image owner accepts the request then a token send to the image user and using this token image user can download the image. After obtaining the image, suppose the image user logged into the system as image owner and tries to upload the same image he or she obtained from another owner won't be a success as

images are watermarked by the owner. The name of such fake user will be available to the admin and admin can block such users. Admin performs watermark extraction to obtain name of fake users.

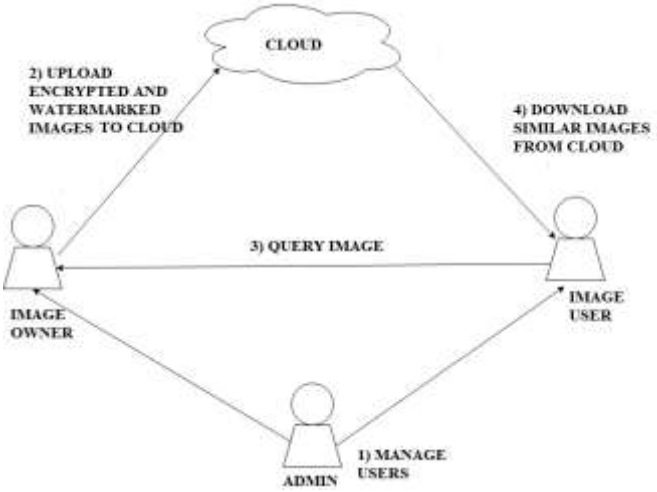

Fig -1: PPCBIR System Model

3. IMPLEMENTATION DETAILS

There are mainly two modules in the system, first admin. An admin can add users to the system or can block users from using system. Second, users, they can be image owner or image users. Image owner can upload images to the cloud server, they encrypt and watermark images before outsourcing. Image owner also performs feature extraction, accepting or rejecting request made by image user for obtaining images. Image owner also sends token to the image user so that they can download image. Image user can search for similar images and can send request to the image owner if similar image is found, they can also download the image if image owner accepts the request. The main job of admin is to find fake users that is a user tries to upload image obtained from another user, he or she won't be able to do so as it is watermarked by the image owner [5] [6] [7]. Admin can obtain the name of such fake users by extracting the watermark from image when user upload an image. Admin can block such fake users.

3.1 Color feature extraction

A pixel is made up of four components ARGB, A for alpha that determines transparency then red , green and blue components that determines the colour of the pixel. Each component is represented using 8 bits so a pixel is 32 bit. Each RGB component has value ranging from 0 to 255. 0 means corresponding colour is absent. 255 means it is fully present. Each pixel is scanned horizontally and vertically and in .net a Class called Color is used to obtain the color of the pixel at each position (x,y). In edge histogram, image is divided into 4*4 blocks. In each block relative frequency of occurrence of 5 types of edges is considered. 5 types of edges are horizontal, vertical, two diagonal edges and non-directional edges.

3.2 Feature from accelerated segment test (FAST)

FAST is a corner point extraction technique. Suppose 'p' is a pixel with intensity 'I' and we have to check whether it is a corner point or not. Let ' t' be a threshold. Consider a circle of 16 pixels around p. p is a corner point if their exists set of ' n 'continuous pixels in circle of 16 pixels that are all either brighter than I+t or darker than I-t. A high speed test was proposed to exclude large number of non-corner points. Only top most, bottommost, leftmost and right most pixels from n continuous set of pixels is considered and tested if it is brighter or darker and p is corner point if atleast three of them are either brighter or darker.

3.3 Least significant bit watermarking (LSBW)

Watermarking means hiding a secret data in the image, secret data can be a cipher text, plain text, audio, video or another image. Images are watermarked so that other users can't illegally distribute the image owned by others. It is used for copyright protection. LSBW is the simplest type of watermarking. In LSB least significant bits of each pixel in the image is replaced with the bits of secret data. It is an invisible watermarking technique which is not visible to human eyes. In this paper image owner name and his or her password is encrypted using AES to create a cipher text. This cipher text is the secret data that is to be hidden inside the image.

3.4 Advanced encryption scheme (AES)

Before outsourcing the images to CS, they are encrypted using AES algorithm. Encryption or Decryption function takes 3 parameters, inputpath: path of the file to be encrypted, outputpath: path where encrypted image to be stored and a key. AES is symmetric key block cipher encryption algorithm. Block ciphers take a number of bits and encrypt them as a single unit. Block cipher performs encryption on block of text and same key is used for encryption and decryption. In AES the original text i.e. clear text is converted into bytes. AES performs all its computations on bytes rather than bits. Hence, AES treats the 128 bits of a plaintext block as 16 bytes. Rfc2898DeriveBytes is an in-built class in the header file using System.Security.Cryptography it takes as input a secret key and bytes to generate a key and initialisation vector (iv). Rfc2898DeriveBytes generates key through call to getbytes() and repeated call to getbytes(32) does'nt generates the same key instead 2 calls of getbytes(32) are appended with a parameter cb whose value is non negative 32,which is equivalent to calling getbytes() once with the parameter 64. Intialization vector(iv) is an arbitary no that can be used along with a secret key for encryption, it prevents dictionary attack. Cryptostream in-built class is used for encryption and

decryption that performs symmetric encryption on streams and is useful to encrypt large amount of data.

3.5 Similar image search

When user provides a query image for searching, it's corner points are generated and stored in a list called search. Similarlily corner points of training sets obtained using FAST algorithm is stored in another list called training. Limit (-5 +5) is applied to all the points in list search. Suppose 250 is a point in search, then on applying limit of (-5+5) we have points 245,246,247,248,249,250,251,252,253,254 and 255.

Points in training list is compared with above values to check whether it is nearest to the above points. Suppose point in training set is 250 then it is considered as a match and counter variable for match is incremented.

Images that have large number of matching points are selected and it's average is calculated to determine the category. The average is found by sum of points matched divided by count of matching points. Whenever an image is searched by user, suppose similar image for horse is searched, then the average corresponding to horse image is incremented.

Table -1: Result obtained for image search based on average ratio

Image Type	Query Image	Average matching ratio
Horse	Horse	9158
Lion		167
Sunset		132

4. APPLICATIONS AND FUTURE SCOPE

Privacy preserving content based image retrieval can be used by doctors to retrieve similar cases of patients and facilitate clinical decision making. It is also used by law enforcement agencies to compare the evidence in the crime scene with evidence stored in database. Some detectives may don't want disclose their evidences with others, so evidences are encrypted before outsourcing. In this paper all the computation is done on the client side which increases client overhead. To overcome this all the computation can be transferred to the server side in future. Cloud server retrieves images similar to a given query image from database of encrypted images without knowing the content. A method based on orthogonal decomposition can be used where overhead is distributed equally likely among client and server.

5. CONCLUSIONS

CBIR retrieves images similar to a query image.

The storage need for such large amount of images is a driving factor for outsourcing services such as cloud storage and computing. However it actually raises new challenges in terms of data privacy. One solution is to encrypt the image database before outsourcing it and run all the computation on client side. Here feature extraction and CBIR service are outsourced to the client side.

REFERENCES

[1] Gudivada, V. N., and Raghavan, V. V., 1995. "Content based image retrieval systems". Computer, 28(9), pp. 18–22.
[2] Hirwane, R., 2012. "Fundamental of content based image retrieval". International Journal of Computer Science and Information Technologies.
[3] Smith, J. R., and Chang, S.-F., 1997. "Visualseek: a fully automated content-based image query system". In Proceedings of the fourth ACM international conference on Multimedia, ACM, pp. 87–98.
[4] Xia, Z., Zhu, Y., Sun, X., Qin, Z., and Ren, K., 2018."Towards privacy-preserving content-based image retrieval in cloud computing". IEEE Transactions on Cloud Computing, 6(1), pp. 276–286.
[5] Hu, S., Wang, Q., Wang, J., Qin, Z., and Ren, K., 2016. "Securing sift: Privacy-preserving outsourcing computation of feature extractions over encrypted image data". IEEE Transactions on Image Processing, 25(7), pp. 3411–3425.
[6] Wang, J. Z., Li, J., andWiederhold, G., 2001. "Simplicity: Semantics-sensitive integrated matching for picture libraries". IEEE Transactions on pattern analysis and machine intelligence, 23(9), pp. 947–963.
[7] Rejito, J., Setiana, D., and Rosadi, R. "Image indexing using color histogram and k-means clustering for optimization cbir".

BIOGRAPHIES

Athira Nair M, She is a post graduation student in college of engineering kidangoor. Specialization in computer and information science.She received the graduation in computer science and engineering from college of engineering kallooppara.

Asha Vijayan, She is an assistant professor in college of engineering kidangoor. Specialization in software engineering. She received graduation in computer science and engineering from SJCEG pala.