Content-Base Image Retrieval in Ecommerce

[1]Bhujbal Shubham, [2]Barve Prashant, [3]Deshmukh Aishwarya, [4]Bora Harshita, [5]Minakshi Vharkate

[1]MIT AOE, Student, School of Computer Science and Information Technology MIT AOE, Alandi, Pune, India

[2]MIT AOE, Student, School of Computer Science and Information Technology MIT AOE, Alandi, Pune, India

[3]MIT AOE, Student, School of Computer Science and Information Technology MIT AOE, Alandi, Pune, India

[4]MIT AOE, Student, School of Computer Science and Information Technology MIT AOE, Alandi, Pune, India

[5]MIT AOE, Professor, School of Computer Science and Information Technology MIT AOE, Alandi, Pune, India

Abstract - Millions of consumers prefer e-purchases over shop purchases, all over the world. So in this case, we are just aiming at efficient retrieval of image from the large and huge database for online shopping. Here CBIR system is focused with the help of that device will be able to learn user's choice of clothing and many other things during e-purchase. When text search is used for item searching by keywords or texts it has some errors in search items like expansion in search and inaccuracy in search result. To overcome this problem image search instead of text search is used. This is just an attempt to help the user to choose the best option of their choice among the huge amount of products and decide exactly what they want quickly by easy search by image retrieval. Different Texture, Color and Shape features to identify relationships in images are applied.

Keywords *- Encrypted data processing, Image Processing, Image Mining*

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 [1,2,9]. 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[10]. 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, it is quite difficult to index image for large database. Secondly, two different people, or even the same person on two different days, may index similar images inconsistently [4,5]. The result of these inefficiencies is a less than optimal search result for the end user of the system. Having a 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 [1,3,7]. 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 [6,9]. 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 [6,7]. Companies that rely on querying thousands of images can't afford to spend hours or even days to find desired images. They clearly desire a system much faster than manual searching, ideally as effective and efficient as, let's say, a relational database. But even database contains large number of records. So it's very tedious to get the exact results. So for these Image retrieval can be very valuable to a big variety of companies' fashion designers may need to find similar textures for a piece of clothing [2,9]. Medicine is potentially the one who can benefit more from CBIR. A hospital can produce daily thousands of images, including x-rays; magnetic resonance images (MRIs), computed tomography (CT) brain scans and biopsies [8,10].

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].

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].

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].

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 generate unique hash codes for various modalities of single instance through collective matrix factorization along with the latent factor model [4].

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].

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].

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].

H. Zhang, J. Yuan, X. Gao, and Z. Chen has 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].

A. Karpathy and L. Fei-Fei proposed a model generating the descriptions of natural language of images and their regions. This approach has 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].

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. EXISTING SYSTEM

Along with the increasing requirements, in recent years, cross-media search tasks have received considerable attention. Since, each modality having different representation methods and correlational structures, a variety of methods studied the problem from the aspect of learning correlations between different modalities. Existing methods proposed to use Canonical Correlation Analysis (CCA), manifolds learning, dual-wing harmoniums, deep auto encoder, and deep Boltzmann machine to approach the task. Due to the efficiency of hashing-based methods, there also exists a rich line of work focusing the problem of mapping multi-modal high-dimensional data to low-dimensional hash codes, such as Latent semantic sparse hashing (LSSH), discriminative coupled dictionary hashing (DCDH), Cross-view Hashing (CVH), and so on.

IV. PROPOSED SYSTEM

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.

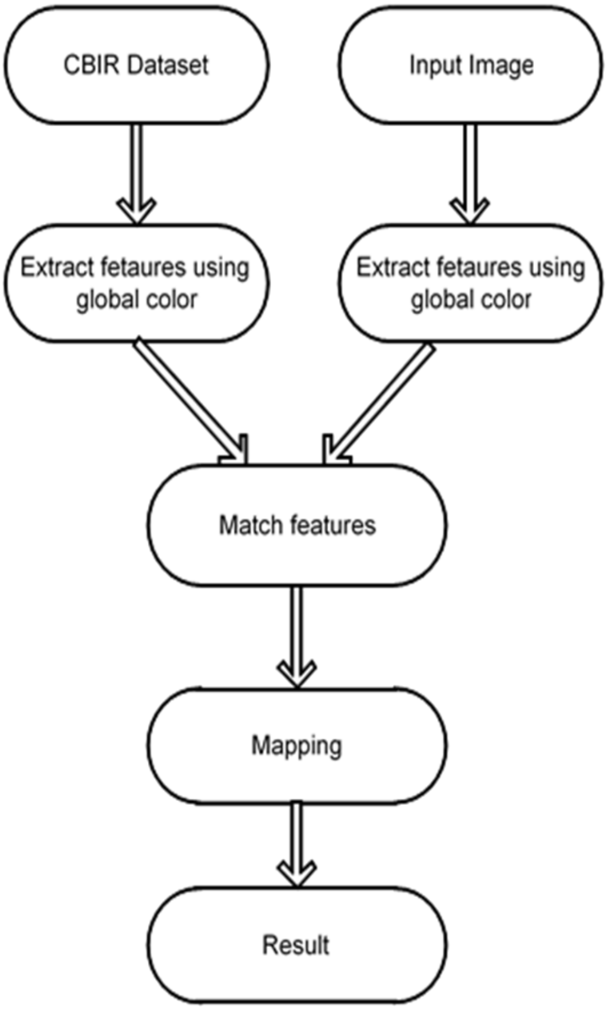

Fig. Flow Diagram

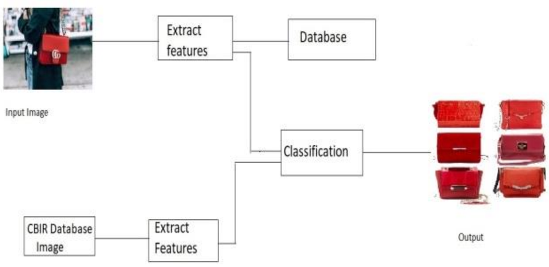

Fig. Retrieval of Image

Fig. Flow Diagram shows the total process from entering input image to retrieving output images. When user inserts input image, in first step AES algorithm is applied for encryption of image. Then feature extraction is done using global color. There are two features are extracted here one is color feature and another is texture feature. For color feature extraction RGB to HSV conversion is used. To extract texture feature Gabor filter is used in this paper. Simultaneously features of database images are also extracted. Then matching of database images with input image is done and similar images are mapped and displayed in the result.

Fig. Retrieval of Image shows the example of image retrieval. User inserts the image of bag, then the process starts with encryption, then both color and texture features are extracted for the bag then similar images from database are also extracted and matched with the bag. For sorting images in database K-means algorithm is used. Finally, images from database are mapped and the best result is displayed.

V. PROPOSED ALGORITHM FLOW

- System Description:
  - ➢ Input: Query Image.
  - ➢ Output: Relevant Images with Query Image
  - ➢ Functions: Color and Texture Feature Extraction
- Color Feature Extraction-

  ○ RGB to HSV Conversion:

  ➢ RGB Values:

$$R' = R/255$$
$$G' = G/255$$
$$B' = B/255$$

$$C_{max} = MAX(R', G', B')$$

$$C_{min} = MIN(R, G', B')$$

$$\Delta = C_{max} - C_{min}$$

  ➢ HSV Values:

$$H = \begin{cases} 60 * (G' - B')/\Delta \ \ mod\ 6, & C_{max} = R' \\ 60 * (B' - R')/\Delta \ \ mod\ 6, & C_{max} = G' \\ 60 * (R' - G')/\ mod\ 6, & C_{max} = B' \\ 0, & \Delta = 0 \end{cases}$$

$$S = \begin{cases} 0, & C_{max} = 0 \\ \Delta / C_{max}, & C_{max} \neq 0 \end{cases}$$

$$V = C_{max}$$

- Texture Feature Extraction-

  ➢ 2-D Gabor filter

$$f(x, y, \omega, \theta, \sigma_x, \sigma_y) = 1/(2\pi\sigma_x\sigma_y) * exp^{[-1/2((x/\sigma_x)^2 + (y/\sigma_y)^2 + j\omega(x\cos\theta + y\sin\theta)]}$$

$$\sigma = \text{Spatial Spread} \quad \omega = Frequency \quad \theta = Orientation$$

  ➢ 1-D Gabor Filter

$$f(x, \omega, \sigma) = 1/(\sqrt{2\pi\sigma})exp^{(-x^2/2\sigma^2 + j\omega x)}$$

  ➢ 1-D Gaussian Function

$$g(x) = 1/(\sqrt{2\pi\sigma_1})exp^{(-x^2/2\sigma^2)}$$

- Success Conditions: Relevant Images are retrieved.

- Failure Conditions: Irrelevant Image or No image is retrieved.

VI. 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.

VII. 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.