A new approach for Content Based Image Retrieval – A Review

Harkamal Kaur[1], Er. Manit Kapoor[2], Dr. Naveen Dhillon[3]

[1,2,3] Department Of ECE, IKPTU, Ramgarhia Institute of Engineering and Technology, Phagwara

---***---

ABSTRACT- *The content based image retrieval (CBIR) methods are used to discover the similar images in accordance with the input image from the database. The image retrieval applications plays a vital role in the case of big databases, where thousands of millions or more images are stored. In the case of media sharing platforms such as Instagram, Whatsapp, Facebook, Picasa, etc, a very large number of data is uploaded on these portals on the daily basis, which makes it impossible to discover the relevant image data manually. Hence there is a strong requirement of versatile information based image retrieval engines from such databases, which can discover the relevant images out of the given database. In this paper, an innovative model for the image retrieval on the basis of color and texture features has been proposed, which is expected to resolve the issue related to the accuracy of image retrieval engines. The performance of the model would be analyzed by using the accuracy metrics such as recall, precision, F1-measure and overall accuracy.*

***Keywords:* CBIR, image processing, visual features, texture features**

INTRODUCTION:

Cloud computing is the conveyance of figuring administrations over the Internet. Cloud administrations enable people and organizations to utilize programming and equipment that are overseen by outsiders at remote areas. Cases of cloud administrations incorporate online record stockpiling, long range informal communication locales, webmail, and online business applications. The cloud computing model enables access to data and PC assets from anyplace that a system association is accessible. Cloud computing gives a mutual pool of assets, including information storage room, systems, PC preparing power, and concentrated corporate and client applications. The qualities of cloud computing incorporate on-request self-administration, expansive system get to, asset pooling, fast flexibility and estimated benefit. On-request self-administration implies those clients (typically associations) can ask for and deal with their own particular processing assets. Expansive system get to enables administrations to be offered over the Internet or private systems. In remote server farms, clients have decision to draw the assets from a pool of processing assets. The quantity of administrations can be either little or vast; and utilization of an administration is estimated and clients are charged as needs be.

The administration models of cloud computing can be delegated: Software as a Service i.e. SaaS, Platform as a Service i.e. PaaS and Infrastructure as a Service i.e. IaaS. In Software as a Service demonstrate, a pre-made application, alongside any required programming, working framework, equipment, and system are given. In PaaS, a working framework, equipment, and system are given, and the client introduces or builds up its own particular programming and applications. The IaaS display gives only the equipment and system; the client introduces or builds up its own particular working frameworks, programming and applications. While there are benefits, there are protection and security concerns as well. Information is going over the Internet and is put away in remote areas. Also, cloud suppliers regularly serve different clients at the same time. The greater part of this may raise the size of presentation to conceivable ruptures, both incidental and ponder. Concerns have been raised by numerous that cloud computing may prompt "capacity crawl" employments of information by cloud suppliers that were not expected when the data was initially gathered and for which assent has regularly not been acquired. Given that it is so economical to keep information, there is minimal motivating force to expel the data from the cloud and more motivations to discover different activities with it.
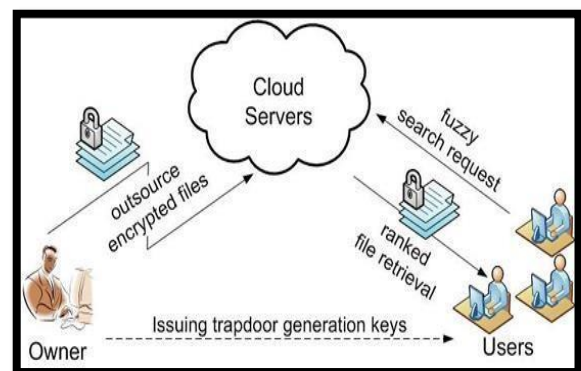

Figure 1: Framework of encrypted cloud data to retrieve the files based on similar search

The need to isolate information when managing suppliers that serve numerous clients, potential auxiliary employments of the information—these are regions that associations should remember while thinking about a cloud supplier and while arranging contracts or inspecting terms of administration with a cloud supplier. Given that the association exchanging this data to the supplier is eventually responsible for its insurance, it needs to guarantee that the individual data is fitting taken care of.

The two elements are broke down as an answer in the current framework which are identified with seek security necessity i.e. watchword security and document secrecy.

- **File secrecy:** Since then, record content must be handled, in this way the quality of the document privacy vigorously relies on security quality of symmetric encryption.

- **Keyword protection:** During the symmetric encryption conspire, the inquiry trapdoor was created so the protection of question watchword relies upon the security quality of the symmetric encryption plot

CBIR TECHNIQUES:

There exist a few methods to recover the pictures however there exist issue of recovering the pictures based on pixels.

Semantic Retrieval: When client influences solicitations to like "discover pictures of Barack Obama" at that point semantic inquiry is begun. Be that as it may, this undertaking is exceptionally hard to perform by PCs. Along these lines bring down level highlights like shading, shape and surface are utilized. The consequences of picture recovery likewise require human input to recognize the larger amount ideas.

Importance Feedback: keeping in mind the end goal to make the utilization of CBIR effective there is have to comprehend the capacity of client purpose. CBIR make utilization of pertinence criticism, where clients stamp the came about pictures as important or not pertinent or

unbiased and afterward supplant the hunt picture with the applicable new data.

Other inquiry techniques: These may incorporate strategies like picture recovery by picture locale, by visual draw, by coordinate particular of picture highlights, by touch, voice and so forth.

Picture Distances Measures: Two pictures can be thought about on premise of their separation measures. Different measurements of pictures are utilized, for example, shading, surface, shape and others. The separation of significant worth 0 demonstrates correct match with picture question. In this manner the outcomes are then put away on premise of their separations to the questioned picture.

Color: Method of picture recovery in this system depends on the measure of shading comparability by registering shading histogram for each picture that connotes the extent of pixels of a picture. This is the most generally utilized method since it can be performed without respect to picture measure. Shading extents are additionally grouped based on area and spatial relationship among a few shading areas.

Surface: This strategy spatially characterizes the picture and furthermore searches for visual examples. Contingent upon the quantity of surfaces distinguished in a picture, they are spoken to as" Texel's" and afterward put into number of sets. This characterizes the area of surface. Surface is recognized by demonstrating it in a two dimensional dark level variety. Strategies to characterize surfaces are co-event network, laws surface vitality and wavelet change.
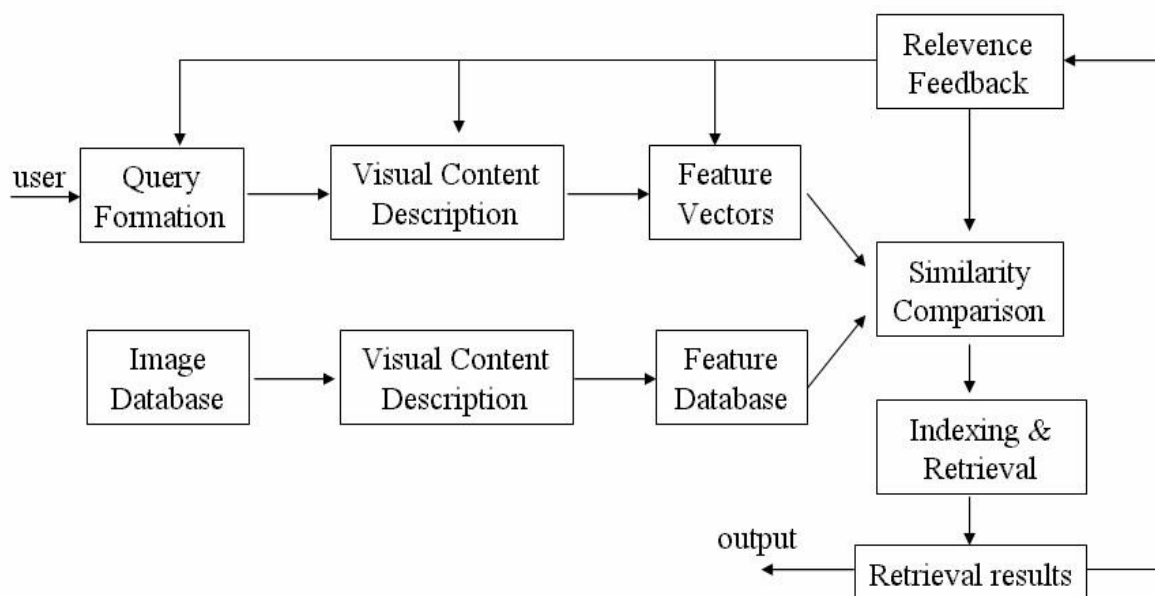

Figure 2: Content based image retrieval

Shape: Shape doesn't think about entire picture yet to state of a specific locale to be searched out. Two procedures are connected first division or edge identification to picture. Shape channels and shape descriptors are additionally utilized. Some shape descriptors incorporate Fourier change and minute invariant

USER FEEDBACK TECHNIQUES FOR CBIR

Significance criticism based intuitive recovery approach considers the two unmistakable attributes of CBIR, first is the hole which exist between the abnormal state ideas and low level highlights of the picture and second is the subjectivity of human impression of visual substance. Accordingly amid the recovery procedure both the qualities are caught by progressively refreshed weights that depend on the client's importance input. At the end of the day we can state that it is utilized to expand precision of the picture being looked. One of the techniques to recover pictures which are utilized to figure the neighborhood include importance is PFRL strategy. The best N comes about are appeared to the end client when some info question picture is given. Presently criticism is required from client side to choose the pictures which are important to the inquiry picture. Here the pictures are characterized in two areas one is containing pertinent pictures and the other is containing disparate pictures, at that point the normal is figured of the two segments. On the off chance that every one of the pictures are disposed of, the arrangement of new pictures is chosen from database. The procedure is proceeded until the point when the client gets the coveted pictures.

Inquiry point development: According to the clients question point ought to be near the curved locale of picture. The space vector equation proposed by Rochhio is given by:

$$Q\_i = Q\_{(i-1)} + \propto \sum\_{(j=1)}^{nr} R\_j/nr - \beta \sum\_{(j=1)}^{nir} 〚IR〛\_j/nir \ (I)$$

Where Q_i is the vector of ith inquiry, R_j is the vector of jth important picture, nr is the cardinality of significant pictures and nir is the cardinality of insignificant pictures

Inquiry Re-weighting: The client takes in question from positive and negative cases by weighting the highlights of picture. The low level visual highlights and abnormal state human ideas are progressively refreshed in the RF approach.

Question Expansion: Query point development and inquiry re-weighting alone can't totally fulfill the client's advantage. To take care of the issue QEX method is utilized and it gives brilliant recovery of pictures. Here the client submit inquiry and as per this question number of pictures which are important are appeared as result.. At

the point when the client found the correct outcomes, at that point there is no want to grow the inquiry more.

Half and half Approach: It is another kind of RF procedure which is once in a while utilized. Here the attention is on the log which is originating from the different clients. The mixture RF method is IRRL. The main issue looked here is one can't maintain a strategic distance from the long cycles of criticism. So this procedure is once in a while utilized.

LITERATURE SURVEY

Song et al. [4] proposed the cryptographic methods for the problem of searching over encrypted data and provided the security proofs for the resulting crypto systems. Techniques have several crucial advantages. They are probably more secure: they provide provable secrecy for encryption, means that the un-trusted server cannot draw anything about the plaintext when only cipher-text is given. Also the un- trusted server cannot learn anything more about the plaintext but only the search result, meaning that they provide query isolation for searches. They provide controlled searching means without the user's authorization, the un-trusted server cannot search for an arbitrary word. They also provide users the facility of hidden queries, so that they may ask the un-trusted server to search for a secret word without revealing that word to the server. Curtmola etal. [2] presented a per-keyword index construction, where each entry of the table represent the whole hash table index which contains the trapdoor for a keyword and an encrypted set of file identifiers. According to this searchable symmetric encryption scheme a party is allowed to outsource the storage of its data to another party in a private manner and maintaining the ability to search over it selectively. Wang et al. [3] proposed that for the first time they formalize and solve the problem of effective fuzzy keyword search over encrypted cloud data as well as maintain the keyword privacy. Fuzzy keyword search is greatly used to enhance system usability by returning only the matching files when users' searching inputs exactly match the predefined keywords or the closest possible matching files based on keyword similarity semantics, when exact match fails. Wang et al. [5] proposed a solution for ranked single-keyword search regarding the certain relevance score. For the first time this paper define and solve the problem of secure ranked keyword search over encrypted cloud data. Ren et al. [6] suggested the similar secure per-file index, where for each file an index including trapdoors of all unique words is constructed. Here, the author proposed several critical security challenges and suggested for future investigation of security solution for a trustworthy public cloud environment. Cao et al. and Yang et al. [1,8] proposed a scheme for multi- keyword ranked search, where inner product similarity is used for result ranking. This paper,

for the first time, defines and solves the challenging problem of privacy preserving multi-keyword ranked search over encrypted cloud data. Xia et al.[7] described that the results could return not only the exactly matched files, but also the files including the terms which are semantically related to the query keyword. Thus in the proposed scheme, a corresponding file metadata is constructed for each file. Now both the encrypted metadata set and file collection are uploaded to the cloud server. With the help of metadata set, the cloud server builds the inverted index and constructs semantic relationship library (SRL) for the keywords set. After receiving a query request, the cloud server first finds out the keywords which are semantically related to the query keyword according to SRL.

PROBLEM FORMULATION

Cloud data retrieval or search is the process of the searching the similar search data against the user query submitted in the form of image or text. Existing search data retrieval algorithm in the base paper supports one keyword queries only. The existing project in the base paper is based on the search method over the encrypted cloud data. One point should be remembered here that the data over CBIR platforms in generally stored in the encrypted form to ensure the data security, which increases the response time. The existing model is based upon the efficient content based image retrieval (CBIR) based on early rejection model and robust feature descriptors for the high performance. The existing model enables the CBIR query search based upon encrypted feature descriptors using the early termination based method. The existing CBIR model can be improved by using the multivariate feature descriptors in the perfect amalgamation to enhance the performance of the existing model. Also the early termination process can be secured using the encapsulated and protected feature comparison in order to reduce the concerned security issues. The one term-at-a-time model can be improved by using the one-image all-terms model to improve the performance.

PROPOSED MODEL

The proposed model will be designed for the multiple features based content based image retrieval from the given database. The proposed model will be based upon the amalgamation of the color and texture features by using multivariate method which will be certainly forming the improved paradigm of the existing content based image retrieval schemes. The system design will be initialized with the acquisition of the query data in the form of the input image by the user, which will further undergo the CBIR model in order to find the similar images as the final result of the CBIR engine. The proposed model will extract thecolor and texture based features from the input query image, which would be matched to

the training data images, whom color and texture based features are already extracted and saved under the training matrix. The feature matching is done in the dual layer design using the probabilistic classification algorithm (support vector machine or SVM) or the non- probabilistic classification algorithm (k-nearest neighbor or KNN), which will produce the similarity based results and ranks the images accordingly. Afterwards the results would be shown on the user's screen as the final results from the final ranking (indexing) matrix, which contains the final results after the dual layer evaluation of the features of the training and testing image data.

CONCLUSION

In this paper a reduced set of similar images is generated by using multivariate method. The proposed multivariate method involves the image color and texture for the purpose of image matching to the query image (also known as a reference image). The most matching entities are returned as the final results by the image extraction method. There are four methods, which involves three singular feature and one multivariate feature based models, have been implemented. The multivariate model has been found much stable and returned the maximum accuracy under this model. The algorithm is similar to the experience the extraction of similar images of human brain, lung cancer by using SVM. Experiment results clearly show the effectiveness of the algorithm.

REFERENCES

[1] Cao, N., Wang, C., Li, M., Ren, K., & Lou, W. (2014), "Privacy-preserving multi-keyword ranked search over encrypted cloud data", Parallel and Cloud Systems, IEEE Transactions , pp. 222-233.

[2] Curtmola, R., Garay, J., Kamara, S., & Ostrovsky, R. (2006, October), "Searchable symmetric encryption: improved definitions and efficient constructions", In Proceedings of the 13th ACM conference on Computer and communications security, pp. 79-88.

[3] Li, J., Wang, Q., Wang, C., Cao, N., Ren, K., & Lou, W. (2010, March), "Fuzzy keyword search over encrypted data in cloud computing", In INFOCOM, 2010 Proceedings IEEE, pp.1-5.

[4] Song, D. X., Wagner, D., & Perrig, A. (2000), "Practical techniques for searches on encrypted data", In Security and Privacy, 2000. S & P2000 Proceeding pp. 44-55.

[5] Wang, C., Cao, N., Li, J., Ren, K., & Lou, W. (2010, June), "Secure ranked keyword search over encrypted cloud data", In Cloud Computing Systems (ICDCS), 2010 IEEE 30th International Conference , pp. 253-262.

[6] Wang, C., Cao, N., Ren, K., & Lou, W. (2012). Enabling secure and efficient ranked keyword search over outsourced cloud data. Parallel and Cloud Systems, IEEE Transactions , pp. 1467-1479.

[7] Xia, Z., Zhu, Y., Sun, X., & Chen, L. (2014). Secure semantic expansion based search over encrypted cloud data supporting similarity ranking. Journal of Cloud Computing, pp. 1-11.

[8] Yang, C., Zhang, W., Xu, J., Xu, J., & Yu, N. (2012, November). A fast privacy-preserving multi- keyword search scheme on cloud data. In Proceedings of the 2012 International Conference on Cloud and Service Computing , pp. 104-110.