# A Robust Image Retrieval Algorithm by Utilizing Encrypted Visual and Textual Features

Soumya Sharma[1], Piyush Jain[2]
[1]*Research Scholar,* [2]*Assistant professor. Dept. of Computer Science, OIST, RGPV, India*

*Abstract*— As the large number of internet users are increasing day by day. So digital data of those users are also increasing, now this increase of data has attracted many researchers. Searching the relevant images from the collection of dataset is an important issue. In this paper two important features are utilize first is textual while second is visual one. For maintaining the privacy of the dataset on server image retrieval process is done on encrypted data. Here model takes two input first one include the textual query while second include visual query. Both query is encrypted first than pass to the servers for comparison with dataset textual and visual features. Encrypted user text query makes initial ranking of the dataset images than Co-Occurrence matrix feature of the visual query makes final ranking of the image. So proposed work provide retrieval of image in high protective environment with less execution time. Experiment was done on real and artificial set of images. Result shows that proposed work is better on different evaluation parameters as compare to previous existing methods.

Keywords— Annotation, Image retrieval, Feature extraction, Re-ranking, Visual contents.

## I. INTRODUCTION

As the internet devices with networking infrastructure is growing day by day. So transferring of data is quit easy for different purposes. Here this increase the data on the servers rapidly, which required a fast an efficient information retrieval system. Although general search engines have been well developed, searching video content over the Web is still a challenging task. Typically, most Web search engines index only the metadata of videos and search through a text-based approach.

As image retrieval is done by two techniques first is text and other is content in case of text based retrieval bag of words are attach with the image. Now that bag of word contains information like keywords, heading, classifying codes, etc [2]. This kind of retrieval is non-standardized as annotation of image is done by a human and as per his language annotation is done. One more issue is that text description is complex and common so image description is done that is totally base on personal views. This kind of image organization for large dataset is impractical where continuous surveillance images are coming one by one [3].

Searching of image retrieval is done by using content of image which is nothing=g except the visual features. As text based retrieval plays an important role in the image mining but to improve the accuracy of fetching image from database, content BIR is done. Here it is a technique which semantically match the image features of the images in the database for clustering, ranking, etc [4]. As human interaction is required to reduce by using text or annotation which is further improve by this CBIR, so this is the main goal of the work. The computer must be able to retrieve images from a database without any human assumption on specific domain (such as texture vs. non-texture, or indoor vs. outdoor).



All Rights Reserved © 2017 IJARCET

Most major task in case of the content base retrieval is defying the rules for the comparison of image features or signatures to other images in the database. As feature comparison includes pixel values. As difference between the features is calculated for finding the relevancy as relevant images have less difference while irrelevant feature has high difference.

The main difficulty in such an approach is to reasonably define the similarity between two images, i.e. to determine if two images show the same content. The authors in [17] calculate the images' distance based on the number of matching local features between two images. This approach works well for landmarks or product images as in such cases typically many images exist showing the exact same object. However, when searching for object categories or scenes it cannot expect to reliably match the local image descriptors. Thus we use a more sophisticated image description based on automatic content analysis. Moreover we do not rely solely on the automatically extracted visual content description for similarity definition, but we also exploit an image description based on the available metadata. More specifically we also use an representation based on the author's tags.

## II. Related Work

Amreen Posharkar et al (March 2015) [1] they proposed a system that capable to serve a hybrid approach to combined CBIR and e-commerce website to retrieval quality products for the user based on internal This system also allowed the user to jot down his/her thoughts since the system enables the user to draw the sketch of the desired product. So ,the result of the search returned will be of good quality demanded by the user. For every search, relevant results will be returned to the user. User’s feedback will be taken into consideration so that more relevant search can be provided. Hence the proposed system provided shopping in a more fun loving manner and at the same time will be more efficient to the user.

Aboli W. Hole et al (March 2015) [2] In this paper, they have focused on recently developed image mining techniques. They explained the basic aim of image mining techniques is to discovering meaningful correlations and formulations from previously collected image data. Many different application areas utilize image mining as a means to achieve effective usage of semantic information about images.

Iyad Aldasouqi and Mahmoud Hassan [4], proposed a fast algorithm for detecting human faces in color images using HSV color model without compromising the speed of detection. The algorithm is fast and can be used in some real-time applications.

Vadivel, A et. al., [5], did a detailed analysis of the properties of the HSV (Hue, Saturation and Intensity Value) color space, laid emphasis on the visual perception of the color of an image pixel with the variation in hue, saturation and intensity values of the pixel. Using the results of this analysis, they determined the relative importance of hue and intensity based on the saturation of a pixel and applied this concept in histogram generation for content-based image retrieval (CBIR) from large databases. In traditional histograms, each pixel contributes only to one component of the histogram. However, they proposed a method using soft decision that contributes to two components of a histogram for each pixel. Shamik

Yogita D. Shinde et al(June 2015) [6] they projected on CBIR algorithm is based on Transform Sectorisation technique and hence it is free from limitations like illumination differentiations among images, rotation variance, needing same sized images used in database. In this total seven variation of Self Mutated Hybrid Wavelet Transform is considered for experimentation. Proposed CBIR technique involves Sectorisation of Self Mutated Hybrid Wavelet Transformed Images for feature extraction. Among all tried combinations, proposed technique gives better performance with SM-Sine transform with 12 sectors used in feature extraction. Performance of proposed algorithm decreases with increase in number of sectors used in feature extraction except SM-Sine Transform. Performance comparison shows that proposed algorithm performs better with only SM-Sine


All Rights Reserved © 2017 IJARCET

transform because for other proposed variations performance is decreased

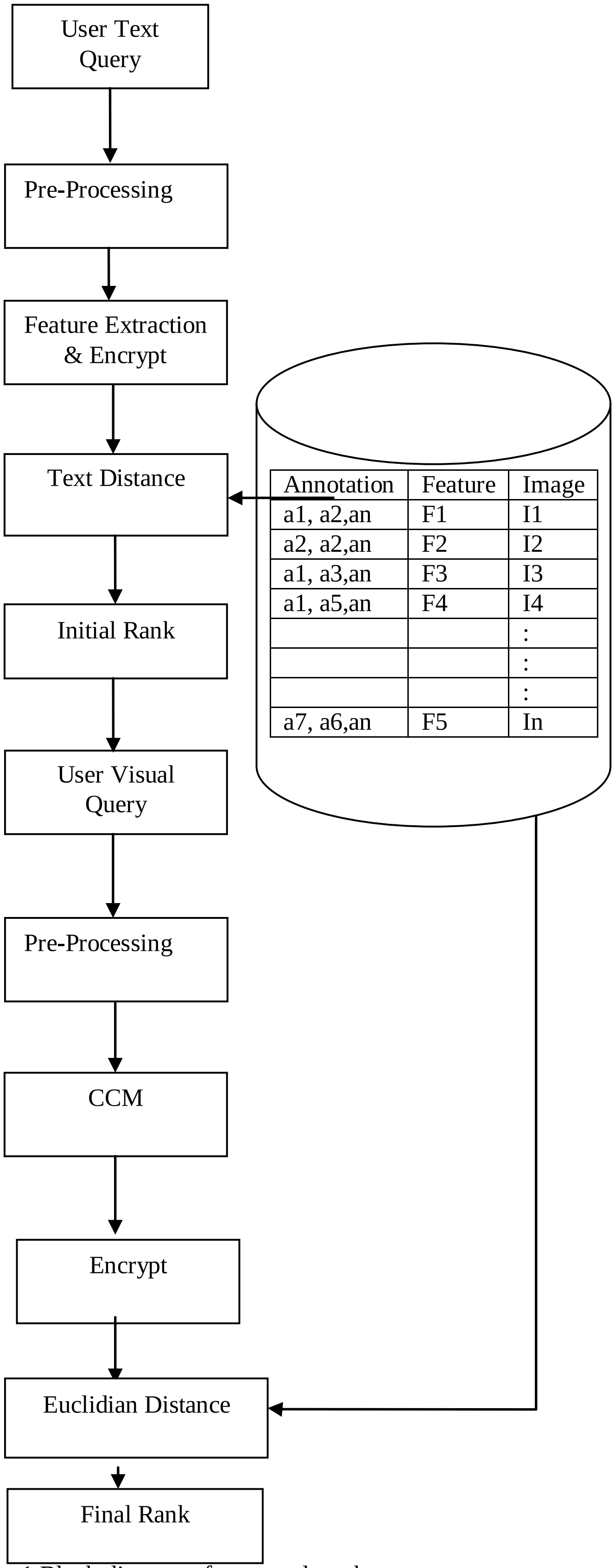


Fig. 1 Block diagram of proposed work.

## III. Proposed Work

**Text Pre-Processing** : Text preprocessing is consisting of words which are responsible for lowering the performance of learning models. Data preprocessing reduces the size of the input text query significantly. Stop-words are functional words which occur frequently in the language of the text (for example a, the, an, of etc. in English language), so that they are not useful for classification.

**Feature Extraction**

Here words which are not stop word are consider as feature of the query. Now assign number to each text of the query. So that a dictionary of words with there number is created where each text is identified by separate number. Such as

D[] = [1, 2, 3, 4, 6, 7, 8, 9]

So for n document has its own vector sequence D[n].

**Pailler Encryption**

This cryptosystem is base on the public and private key concept. Here input vector D[n], will be encrypt by this algorithm.

1. Choose two large prime numbers $p$ and $q$ randomly and independently of each other such that gcd($p$q,($p$−1)($q$−1))=1.

2. Compute RSA modulus $n = pq$ and Carmichael's function $\lambda\lambda = llllll(pp-1, qq-1)$
3. $\mu$=mod(n) / ($L$($g\lambda$ mod($n$*n)−1

   Where the function $L$ is defined as $L(u)=(u-1)/n$ .

So The public key is ($n$, $g$), private key is (λ, μ).

**Text Distance**

Here input query after encryption is transform in other numeric value. So conversion of same text have same value


All Rights Reserved © 2017 IJARCET

for comparison. This can be understand as let “College” word have numeric value 28 after encryption its transform value is 2456. So if “college” present at server for image keyword then its transform value is also 2456 only for same set of encryption key.

In this step count of similar query words found in image keywords is use for ranking. This can be understand as let query be {2456, 1324, 2783} and I1 content is {2456, 1324, 2711}, while I2 content is {1256, 1114, 2783} then distance of query from I1 and I2 is [2, 1]. Base on distance vector I1 image has high rank as compare to I2.

**Visual Pre-Processing**

In this step image is resize in fix dimension. As different image have different dimension. So conversion of each is done in this step. One more work is to convert all image in gray format. AS different image has RGB, HSV, etc. format so working on single format is required.

**Co-occurrence Matrix (CCM)**

In order to get the texture of the image one of the important method is CCM. Here CCM present th texture property by the correlation of the neighbouring pixels [5]. It quantificational describes the texture feature. In this paper four features is selected including energy, contrast, entropy, inverse difference.

$$InverseDifference = \sum_{i=1} \sum_{j=1} \frac{1}{(1+(i-j)^2} m(i, j)$$

where m(i, j) the value in cell (i, j). The properties of an image texture are detected indirectly by using the co-occurrence matrix from which special indexes called “image indicators” are exploited. The gray level co-occurrence matrix (CCM [7]) is a feature that evaluate different values for the texture comparison between images.. The indicators calculated in this work are:

$$Entropy = -\sum_{i=1} \sum_{j=1} m(i, j) \log[m(i, j)]$$

The Entropy indicator measures the disorder or complexity of an image. The highest value of Entropy is found when the values are allocated quite uniformly throughout the matrix. This happens when the image has no pairs of grey level, with particular preference over others. Entropy is strongly but inversely correlated to Energy.

$$Energy = \sum_{i=1} \sum_{j=1} (m(i, j))^2$$

This statistic measures, the textural uniformity, it detects disorders in textures. This parameter indicates how much the texture is homogeneous, i.e. the CCM contains values distributed fairly uniformly over all grids. It is high when the CCM has few entries of large magnitude, low when all entries are almost equal. This is a measure of local homogeneity.

$$Contrast = \sum_{i=1} \sum_{j=1} (i-j)^2 * m(i, j)$$

Texture based retrieval has been popular particularly when surfaces with repetitive patterns are involved. Different methods have been used to express the basic properties such as the granularity, gray-level distribution and directionality of the pattern.

**Visual Distance**

This can be understand as Let X be a query image matrix and Y be the dataset image matrix. Then distance between them is calculate by:

$$D = \sqrt{sum((X - Y)^2)}$$



All Rights Reserved © 2017 IJARCET

Base on the minimum distance value between query and dataset image rank is assigned to the image. This is consider as final rank of the work.

**Proposed Work Image Retrieval**

Input: Text_Query (TQ), Visual_Query(VQ)

Output: Image_Rank(IR)

step 1. Keyword <- Preprocessing(TQ)

step 2. TS <- Text_Encryption(Keyword)

step 3. TS <- Text_Similarity(Keyword)//TS: Image text similarity

step 4. TR<- Initial_Rank(TS) // Text base initial Rank

step 5. CQ←CCM_feature(VQ)

step 6. Loop i=1:m // m: top m images in the TR vector

step 7. D[m]←Eludician_Dist(CQ, C, EQ, E) // Find distance from query image

step 8. end Loop

step 9. IR←Sort_asending(D) // FR: Final Rank

## IV. Experiment And Result

In this section, first introduce experimental settings, and then present the experimental results that validate the effectiveness of the approach. The experiments actually contain two parts. This work is compare with other several existing methods that adopt all features.

**Evaluation Parameter**: NDCG [6, 12] as the performance evaluation measure.

The NDCG measure is computed as

$$NDCG@P = Z_P \sum_{i=1}^{P} \frac{2^{l(i)} - 1}{\log(i+1)}$$

where $P$ is the considered depth, $l(i)$ is the relevance level of the $i$ -th image and $ZP$ is a normalization constant that is chosen to let the optimal ranking's NDCG score to be 1.

**Data Sets**

In order to conduct the experiment an artificial dataset which is a collection of images from different category are utilize. Now this collection of images of different category are shown in table 1 for which one can make some important keyword collection for different images.

| Category | Examples |
|---|---|
| Objects | Ipod, map |
| Insect | Butterfly, Gorrila |
| Scene | Taj Mahal, Hotel Taj |
| Human | Barack Obama, Lena |

Table 1 Dataset of Different category.

Results

| Comparison of Proposed and Previous work Execution Time | | |
|---|---|---|
| Environment | Proposed Work | Previous Work [] |
| Objects | 4.245 | 9.203 |
| Animal | 4.070945 | 9.134815 |
| Scene | 3.705 | 10.3299 |
| Person | 4.572 | 8.025 |

Table 5.2 Comparison of proposed work and previous work on execution time in seconds.


All Rights Reserved © 2017 IJARCET

From the above table it is find that the including of the CCM feature for query has increase the efficiency of image re-ranking. So execution time value of various environment images are lower as compare to most of previous approaches.

| Comparison of Proposed and Previous work NDCG @10 | | |
|---|---|---|
| Environment | Proposed Work | Previous Work [] |
| Objects | 0.648932 | 0.496087 |
| Animal | 0.800694 | 0.330689 |
| Scene | 0.936379 | 0.251903 |
| Person | 0.644762 | 0.286346 |

Table 5.3 Comparison of proposed work and previous work on NDCG @7.

From the above table it is find that the including of the Text feature for query has increase the efficiency of image re-ranking. So NDCG value of various environment images are higher as compare to most of previous approaches.

| Comparison of Proposed and Previous work NDCG @7 | | |
|---|---|---|
| Environment | Proposed Work | Previous Work [] |
| Objects | 0.810462 | 0.576215 |
| Animal | 1 | 0.369646 |
| Scene | 1 | 0.274876 |
| Person | 0.637438 | 0.274876 |

Table 5.4 Comparison of proposed work and previous work on NDCG @7.

From the above table it is find that the including of the Text feature for query has increase the efficiency of image re-ranking. So NDCG value of various environment images are higher as compare to most of previous approaches.

## V. Conclusions

Web image re-ranking has been widely used to reduce the user searching time on the internet; its success mainly depend on the accuracy of image features similarities. This work present utilizing of the new text as well as visual features for ranking the image as both make the re-ranking process more powerful, which is shown in results. In future in order to improve the efficiency more features of images will be include. As to reduce the server time of making graph and re-ranking one filter need to be inserted into this so that it will filter relevant and irrelevent images at the initial stage.

## VI. References


1. Aboli W. Hole and Prabhakar L. Ramteke, Design and Implementation of Content Based Image Retrieval Using Data Mining and Image Processing Techniques, International Journal of Advance Research in Computer Science and Management Studies ,Volume 3, Issue 3, March 2015 pg. 219-224.
2. Supreet Sidhu and Jyoti Saxena, Content Based Image Retrieval A Review, International Journal Of Research In Computer Applications And Robotics, Vol.3 Issue.5, Pg.: 84-88 ,May 2015.
3. Sridhar, G., "Color and Texture Based Image Retrieval, (2009).
4. I. Aldasouqi and M. Hassan, "Human face detection system using HSV", Proc. Of 9th WSEAS Int. Conf. on Circuits, Systems, Electronics, Control & Signal Processing (CSECS'10), Atenas, Grecia, (2010).
5. A. Vadivel, A. K. Majumdar and S. Sural, "Perceptually smooth histogram generation from the HSV color space for content based image retrieval", International Conference on Advances in Pattern Recognition, (2003).
6. Yogita D. Shinde and Sudeep D. Thepade, Content based Image Retrieval using Sectorisation of Self Mutated Hybrid Wavelet Transforms, International Journal of Computer Applications (0975 – 8887) Volume 120 – No.24, June 2015.
7. Dahane, G. M., & Vishwakarma, S., "Content Based Image Retrieval System", IJEIT, Vol. 1, (2012), Pp. 92-96.
8. Chun, Y. D., N. C. Kim, And I. H. Jang, "Content-Based Image Retrieval Using Multiresolution Color And Texture Features", Multimedia, IEEE Transactions On, Vol. 10, No. 6 (2008), Pp. 1073-1084.
9. Chun, Y. D., S. Y. Seo, And N. C. Kim, "Image Retrieval Using BDIP And BVLC Moments", Circuits And Systems For Video Technology, IEEE Transactions On, Vol.13, No. 9, (2003), Pp.951-957.
10. Huang, Y-L., K-L. Wang, and D-R. Chen, "Diagnosis of breast tumors with ultrasonic texture analysis using support vector




All Rights Reserved © 2017 IJARCET

machines", Neural Computing & Applications, vol. 15, no. 2 (2006), pp. 164-169.

11. Ju, Y. A., "Face recognition using local statistics of gradients and correlations", Proc. European Signal Processing Conf., (2010).
12. Meng Wang, Hao Li, Dacheng Tao, Ke Lu, and Xindong Wu “Multimodal Graph-Based Reranking for Web Image Search. IEEE Transaction on image processing Vol. 21, NO. 11, November 2012.
13. G. Bradski and A. Kaehler, Learning OpenCV : Computer Vision with the OpenCV Library, O'Reilly, Sebastopol, CA, 2008.
14. Herbert Bay, Andreas Ess, Tinne Tuytelaars, Luc Van Gool, SURF: Speeded Up Robust Features", Computer Vision and Image Understanding (CVIU), Vol. 110, No. 3, pp. 346--359, 2008.



All Rights Reserved © 2017 IJARCET