An Enhancement of CBIR using LTrP with Hadoop

Pooja R. Gurav [1], Vijaysinh. G. Chavan [2], Abhijit A. Rajguru[3]
[1,2,3] *Department of Computer Science & Engg, Solapur University, Solapur*
SKN Sinhgad College of Engineering, Korti, Solapur, MS, India
[1]poojagurav5@gmail.com, [2]vijay.chavan2008@gamail.com, [3]abhijitcse08@gmail.com

Abstract - Today huge quantity of data or information in the world is in the form of images and which is produced through digital cameras, mobile phones and photo editing software. A large part of this data is stored in online database. This data is private to each user and consequently it should not be accessible by admin or others. Here, a 'content based' search has been proposed which analyzes the content of the image. CBIR System will allow user to upload image and depending on combined values of color, shape and texture, the system will retrieve similar images from database. We propose a novel image indexing and retrieval algorithm using local tetra patterns (LTrP) for content-based image retrieval (CBIR).

Keywords - CBIR, LTrP, Hadoop, Feature Vector, Retrieval.

I. INTRODUCTION

Nowadays, technology has become progressively advanced in terms of security. This has led to cheap and advanced multimedia devices which have improved huge data volumes. This large data needs to be stored in database for various applications such as medical, security, social sites etc. These applications have created a requirement for effective and efficient methods of storage, search and retrieval of images via parallel processing with content related searching techniques with more security. With growing technology, security has become a crucial concern for visual forms of information. For example, suppose an organization has introduced an application for their use. The organization will wish that the input images as well as images in their database are not exposed publicly by the algorithm or any other means such as database admin. We attempt to find a method for storing and retrieving of images in such way that our objective is achieved. Earlier, the images were stored with related labels, names and searching was done on the basis of those labels of paths. But, such a method leads to many errors. If the images are wrongly annotated, named, irrelevant images might be retrieved. Moreover, it is a difficult task to search for an image which has been assigned a wrong label [3].

Queries of image are resolved by computing the same set of features from the input image and the images with most similar features are given as output from the database [6].

A. Content Based Image Retrieval (CBIR)

The term "Content-Based Image Retrieval" is used for retrieving the corresponding images from the database based on their feature of images which derived the image itself like texture, color and shape and domain specific like human faces and fingerprints. The retrieval based on the content of an image is becomes more effective than the text based which is called content based image retrieval which is used for a various applications [2].

1) ***Color:*** Color is the most commonly used attribute or feature in image retrieval systems. Retrieval in these systems is done on the basis of similarity of color. For each image in database, color histogram is computed on each pixel which shows pixel position of each color in the image & their values [4]. Most commonly used color based image retrieval methods are RGB and HSV.

2) ***Texture:*** Texture feature matching may not seem to be useful for image retrieval. Texel's i.e. intensity of pixel is made of intensity of pixel. They help in differentiating between textured and non-textured images from database. The calculations of Texel's may be defined based on the direction of texture, then its coarseness and regularity. Texture can be classified like fine, gross with respect to pixel of an image. And this helps to retrieve textured regions in images on the basis of similarity to automatically-derived towards representing important classes of texture within the collection of images [5].

3) ***Shape:*** At a primary level, shape is one of the basic attributes or feature of any image. Image retrieval methods also use shape. A number of features are computed by the shape of the objects in image. Features are defined at two levels - local and global. These features include aspect ratio, circularity and moment invariant, segments of consecutive boundary, etc. Shape matching does not result accurately in case of 3D objects in images.

II. LITERATURE REVIEW

This paper proposed a novel image indexing and retrieval algorithm using local tetra patterns (LTrPs) for content-based image retrieval (CBIR). The standard local binary pattern (LBP) and local ternary pattern (LTP) encode the relationship between the referenced pixel and its surrounding neighbors by computing gray-level difference. The proposed method encodes the relationship between the referenced pixel and its neighbors, based on the directions that are calculated using the first-order derivatives in

vertical and horizontal directions. The performance of the proposed method is compared with the LBP, the local derivative patterns, and the LTP based on the results obtained using benchmark image databases. [1] First, HSV color space is quantified rationally. Color histogram and texture features based on a co-occurrence matrix are extracted to form feature vectors. Then the characteristics of the global color histogram, local color histogram and texture features are compared and analyzed for CBIR. Based on these works, a CBIR system is designed using color and texture fused features by constructing weights of feature vectors. The relevant retrieval experiments show that the fused features retrieval brings better visual feeling than the single feature retrieval, which means better retrieval results. [2] The content based image retrieval is an interesting and complex problem studied by many researchers all over the world. The complexity is due to retrieval accuracy and retrieval time. From literature survey it is concluded that a wide variety of CBIR method have been proposed in different papers.

III. EXISTING SYSTEM

Earlier, the images were stored with associated labels, names and searching was done on the basis of those labels, names. But, such a method leads to many errors. If the images are wrongly annotated or named so that irrelevant images might be retrieved. Moreover, it becomes a difficult task to search for an image which has been assigned a wrong label [3]. Traditionally, search of the images are using text, tags or keywords or annotation or name assigned to the image while storing into the databases. Whereas if the image which is stored in the database are not uniquely or specifically tagged or wrongly described then it's insufficient, laborious, difficult and extremely time consuming job for search the particular image in the large set of databases [4].

IV. PROPOSED SYSTEM

From the concept of Image processing the proposed system have the following phases or modules:

1. Images Upload
2. Calculate feature Vector (i.e. on Colour & Texture)
3. Compare
4. Search

Where for Colour feature we can use Colour Histogram & for Texture feature we proposed to use Local tetra Patterns (LTrP).

V. SYSTEM ARCHITECTURE

The proposed system incorporates following phase's 1) Upload the images and 2) Search the image from the database. Both modules are deployed on Hadoop cluster. For uploading phase, the data is given as input for storage of HDFS and Query Image is searched from the database.

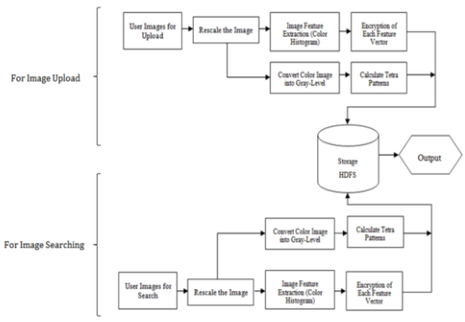

Fig. 1 System Architecture

I. METHODOLOGY

The proposed system incorporates following phases. Those modules are deployed on Hadoop cluster. For uploading module, the data is given as input for storage of HDFS and Query Image is searched from the database.

1) **Upload Images:** System allows user to upload one or more images at a single point into the system. In this proposed system, Upload image module has sub processes. The user uploads the images on database of Hadoop which is called HDFS.

   a) **Phase 1:** User can upload one or more images into the database by giving the image path at a single point from the GUI. The uploaded images by the user may be of different sizes. Large or different sized images may take more time for computation. To deal this problem, the system will have the facility of rescaling the image. After the image has been rescaled, it is converted into gray scale. The first-order derivatives of vertical and horizontal axes are applied and every direction of pixel that is calculated based on the center pixel forms patterns which are divided into four parts. Then tetra patterns are calculated and are separated into three binary patterns. Magnitude of center pixel is computed. The feature extracted from the binary patterns and binary patterns obtained from the magnitude patterns are combined to form a feature vector. The converted image is then stored in .text file format on database.[1]

   b) **Phase 2:** After the image is rescaled, as described in phase 1, its color features are calculated i.e. RGB values for each pixel of image. These calculated feature values are encrypted with simple encryption and a .txt file is formed from the result which is then stored into the database.

2) **Searching of Images:** Like uploading phase, proposed system has search module which allows user to search and retrieve the images from the database by the query image. When a user wish to search a particular image. The searched image is matched with the help of pixel value for LTrP method. It is also searched by color

method where the color encrypted values are compared with the values stored in database. Finally we can compare the results of Image search by color & image search by LTrP method. So that we can easily analyze the implemented technique over another technique.

3) **Image Database Privacy:** In this system, the image stored is in the encrypted values into .text format on the HDFS database of hadoop. This format, in raw form, does not provide any information about the image.

4) **Local Patterns:**

a) LBPs

The LBP operator was introduced in [7] for texture classification. Given a centre pixel in the image, the LBP value is computed by comparing its gray value with its neighbours, as shown in Figure, based on where is the gray value of the centre pixel, is the gray value of its neighbours, is the number of neighbours, and is the radius of the neighbourhood.

$$LBP_{P,R} = \sum_{P=1}^{P} 2^{(P-1)} * f_1(g_p - g_c)$$

$$f_1(x) = \begin{cases} 1, & x \geq 0 \\ 0, & else \end{cases}$$

b) LTPs

The LBP to a three-valued code called the LTP, in which gray values in the zone of width around are quantized to zero, those above are quantized to 1, and those below are quantized to 1, i.e., indicator is replaced with three-valued function [3] and the binary LBP code is replaced by a ternary LTP code, as shown in Fig. 1, i.e.,

$$f_{1(x,g_c,t)} = \begin{cases} +1 & x \geq g_c + t \\ 0 & |x - g_c| < t \\ -1 & x \leq g_c - t \end{cases}$$

c) LDPs

The LDPs for face recognition[7]. They considered that LBP as the non-directional first-order local pattern operator and extended it to higher order (nth-order) called the LDP. The LDP contains more detailed discriminative features as compared with the LBP.

$$LDP_{\alpha}^{n}(g_c) = \sum_{p=1}^{p} 2^{(p-1)} * f_2(I_{\alpha}^{(n-1)}(g_c), I_{\alpha}^{(n-1)}(g_p))|_{p=8}$$

$$f_2(x,y) = \begin{cases} 1, & if\ x.y \leq 0 \\ 0, & else \end{cases}$$

d) LTrP

Local Tetra Patterns are the combination of LBP, LDP and LTP. It describes the special local texture structure using centre gray pixel by direction. The centre level direction of gray level pixel is denoted by I. denotes the centre pixel in I, h horizontal and vertical neighbourhoods of respectively. Each part is converted into three binary patterns by tetra patterns. [1]

$$I_0^1(g_c) = I(g_h) - I(g_c)$$
$$I_{90^\circ}^1(g_c) = I(g_v) - I(g_c)$$

$$M_{I^1(g_P)} = \sqrt{(I_{0^\circ}^1(g_P))^2 + (I_{90^\circ}^1(g_P))^2}$$

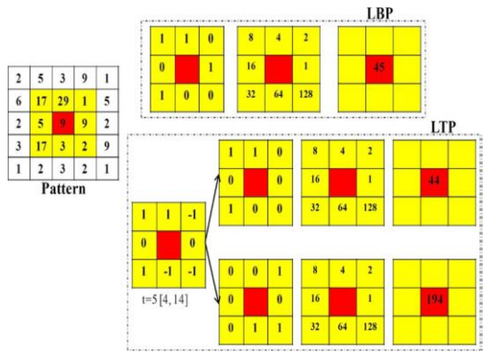

Fig. 2 A LBP & LDP calculation about a pixel

VI. ALGORITHM

Algorithm 1: Upload Image by LTrP Pattern

1. Rescale the image.
2. Convert it into gray scale
3. Apply the calculation of the direction of every pixel.
4. Divide the pattern into four parts based on the direction of the center pixel.
5. Calculate the tetra patterns and separate them into three binary patterns.
6. Save into Database.

Algorithm 2: Search Image by LTrP Pattern

1. Upload the image for search from user.
2. Rescale the image.
3. Convert it into gray scale
4. Apply the calculation of the direction of every.
5. Divide the pattern into four parts based on the direction of the center pixel.
6. Calculate the tetra patterns and Separate them into three binary patterns.
7. Compare the query image with the images in the Database.
8. Result.

Algorithm 3: Upload image by color

1. Rescale the image.
2. Calculate Image feature Extraction (Color Histogram (RGB)) values of each Image.
3. Encryption of color values.
4. Save into database.

Algorithm 4: Search Image by Color

1. Rescale the image.
2. Calculate Image feature Extraction (Color Histogram (RGB)) values of each Image.
3. Encrypt these values.
4. Compare the query image with the Database on the encrypted color values.
5. Result.

VII. CONCLUSION

As we know the images are growing through the various digital devices and these images are added into the image databases and internet for various applications. These images need to be stored and retrieved in effective and efficient manner, because the privacy is the main issue. So that here we have proposed a novel approach referred as LTrPs for CBIR. Where the LTrP encodes the images based on the direction of pixels that are calculated by horizontal and vertical derivatives. [1] For more security also proposed to implement an encryption technique to database on HDFS. Because of effectiveness of this technique it can be suitable for other patterns recognition applications such as face recognition, fingerprint recognition etc.

REFERENCES

[1] Subrahmanyam Murala, R. P. Maheshwari, Local Tetra Patterns: A New Feature Descriptor For Content-Based Image Retrieval, IEEE Transactions On Image Processing, Vol. 21, No. 5, May 2012.

[2] Jun Yue, Content-based image retrieval using color and texture fused features, ELSVIER, 2011.

[3] Veena I Patil , Supriya Kotyal, Survey On Content Based Image Retrieval Using Color And Texture Features, International Conference On Recent Trends In Engineering Science And Management ,Jawaharlal Nehru University, Convention Center, New Delhi (India), 15 March 2015 .

[4] Kommineni Jenni: Content Based Image Retrieval Using Color Strings Comparison, 2nd International Symposium on Big Data and Cloud Computing (ISBCC'15), 2015.

[5] Hinge Smita, Gaikwad Monika : Content Based Image Retrieval Using HADOOP MapReduce, International Journal of Computer Science Trends and Technology (IJCST) – Volume 2 Issue 6, Nov-Dec 2014.

[6] Fatemeh Alamdar: A New Color Feature Extraction Method Based on Quad Histogram, 2011 3rd International Conference on Environmental Science and Information Application Technology (ESIAT 2011).

[7] Ms. Shraddha D: An Image Indexing and Retrieval Method Using Local Tetra Pattern for Content-Based Image Retrieval (CBIR), International Journal of Computer Science and Information Technologies, Vol. 6 (3), 2015.

[8] Ashish Oberoi, Varun Baksh1:A Framework for Medical Image Retrieval Using Local Tetra Patterns, International Journal of Engineering and Technology (IJET), Vol. 5 No 1 Feb- Mar 2013

[9] T. Anucha and P. Wichian, On-line Content-Based Image Retrieval System using Joint Querying and Relevance Feedback Scheme, WSEAS Transaction on Computers, Issue 5, Volume 9, May 2010.

[10] Ji-quan ma: Content-Based Image Retrieval with HSV color space and Texture Feature, International conference on web information systems and mining, IEEE (2009).

[11] Prof.Deepti Chikmurge: *Implementation of CBIR Using MapReduce Over HADOOP*, International Journal of Computer, Information Technology & Bioinformatics (IJCITB), ISSN: 2278-7593, Volume-2, Issue-2.

[12] Swapnil P. Dravyakar et al *"Private Content Based Multimedia Information Retrieval Using Map-Reduce",* International Journal of Computer Science Engineering and Technology (IJCSET), April 2014, Vol 4, Issue 4,125-128, ISSN-2231-0711.

[13] J. Shashank: Content Based Image Retrieval Using Color, Texture and Shape Features, *International Conference on Advanced Computing and Communications, ADCOM* , pp.780 – 784,2007.
Khan, S.M.H, Hussain, A.and Alshaikhli, I.F.T: Comparative Study on Content-Based Image Retrieval (CBIR), *International Conference on Advanced Computer Science Applications and Technologies* (ACSAT), 2012.

[14] Liangliang Shi , Bin Wu ,Bai Wang and Xuguang Yan Map/reduce in CBIR application," *International Conference on Computer Science and Network Technology (ICCSNT),* Vol.4 , pp. 2465 – 2468, Dec. 2011.