

## Image Retrieval Using Content Information

Tiejun Wang, Weilan Wang

School of mathematics and computer science institute, Northwest University for Nationalities,  
Lanzhou 730030, China

Tel.: +86 0931-7663361, fax: +86 0931-4956109  
E-mail: wtj@mail.lzjtu.cn

Received: 14 August 2014 /Accepted: 20 October 2014 /Published: 31 December 2014

**Abstract:** With the rapid development of the multimedia technologies such as compression and digital image sensor technologies, it is more and more necessary to build up a large database. Therefore robust and efficient methodology for image retrieval techniques has gained the attention of researchers. In the last few years the booming interest in the web images has paved a way to make people on content based Image retrieval techniques and also supporting systems for the retrieval process. In this paper, a general review of Content based image retrieval process by comparing with existing retrieval techniques along with experimental results and applications are presented. Some of the major challenges in the adaptation of content based image retrieval algorithms are also discussed. Copyright © 2014 IFSA Publishing, S. L.

**Keywords:** Image retrieval, Annotation, Image analysis, Indexing of images.

### 1. Introduction to CBIR

Image processing is growing field nowadays. It has its own applications and can be applied to many fields. Nowadays, images play a prominent role. It has an impact in many vast fields like satellite, medical, commercial, education, etc. In many areas of commerce, government, academia, and hospitals, large collections of digital images are being created. Many of these collections are the product of digitizing existing collections of analogue photographs, diagrams, drawings, paintings, and prints. Usually, the only way of searching these collections was by keyword indexing, or simply by browsing. The best example is searching for an image on internet by giving keyword.

In fact, different users tend to use different words to describe a same image characteristic. The lack of systematization in the annotation process decreases the performance of the keyword-based image search. These shortcomings have been addressed by the so-

called Content-Based Image Retrieval (CBIR) systems. In these systems, image processing algorithms (usually automatic) are used to extract feature vectors that represent image properties such as color, texture, and shape. In this approach, it is possible to retrieve images similar to one chosen by the user (query-by-example). One of the main advantages of this approach is the possibility of an automatic retrieval process, contrasting to the effort needed to annotate images. Here we brief about the survey of technical aspects of current content-based image retrieval systems.

### 2. Literature Survey

The research efforts in the area of CBIR started since 1980's along with the discussion about image retrieval technologies used in the current commercial image search engines. Three clusters of image retrieval technologies are described: text-based image

retrieval (TBIR), content-based image retrieval (CBIR) and web-based image retrieval (WBIR). By analyzing the advantages and shortcomings of various image retrieval technologies, people discover the ideas for improving the performance of current image search systems by combining WBIR and CBIR methods.

In recent years, there is a rapid increase in the online image collections. According to a study in 1999, there are over 180 million images on the Web, and a million more images are produced every day. However, this huge amount of information cannot be used unless they can be searched and retrieved efficiently. Therefore, image retrieval has been a very active research topic in the recent years. Different users may need access to images based on different image features, including high-level features, such as abstract concepts and keywords, and low-level features, such as color, texture and shape.

Texture analysis has been extensively used in computer vision and pattern recognition applications due to its potential in extracting the prominent features [1]. Texture retrieval is a branch of texture analysis that has attracted wide attention from industries since this is well suited for the identification of products such as ceramic tiles, marble, parquet slabs, etc. The application of the DWT using generalized Gaussian density with Kullback-Leibler distance has shown to provide efficient results for texture image retrieval [3] and image segmentation [2]. However, the DWT can extract only three directional (horizontal, vertical, and diagonal) information from an image. To address this directional limitation, Gabor transform (GT), rotated wavelet filters [4] have been proposed for texture image retrieval. The local binary pattern (LBP) feature has emerged as a silver lining in the field of texture classification and retrieval. Ojala et al. proposed LBPs [5], which are converted to a rotational invariant version for texture classification [6, 7]. Various extensions of the LBP, such as LBP variance with global matching [8], dominant LBPs [9], completed LBPs [10], joint distribution of local patterns with Gaussian mixtures [11], etc., are proposed for rotational invariant texture classification.

A novel image indexing and retrieval algorithm using local binary patterns (LBPs), local ternary patterns (LTPs) and local tetra patterns (LTrPs) for content-based image retrieval (CBIR) have increased the performance [12-17]. 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 newly 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. In addition, we propose a generic strategy to compute nth-order LTrP using (n-1) th-order horizontal and vertical derivatives for efficient CBIR and analyze the effectiveness of our

proposed algorithm by combining it with the Gabor transform.

### 3. New Ideas and Approaches

We do not have yet the considerable and universally acceptable algorithmic means of human vision, more specifically towards image understanding. Therefore, it is not surprising to see various continuing efforts towards the retrieval of images. Some of the key aspects of content based image retrieval and annotation are discussed in the following section.

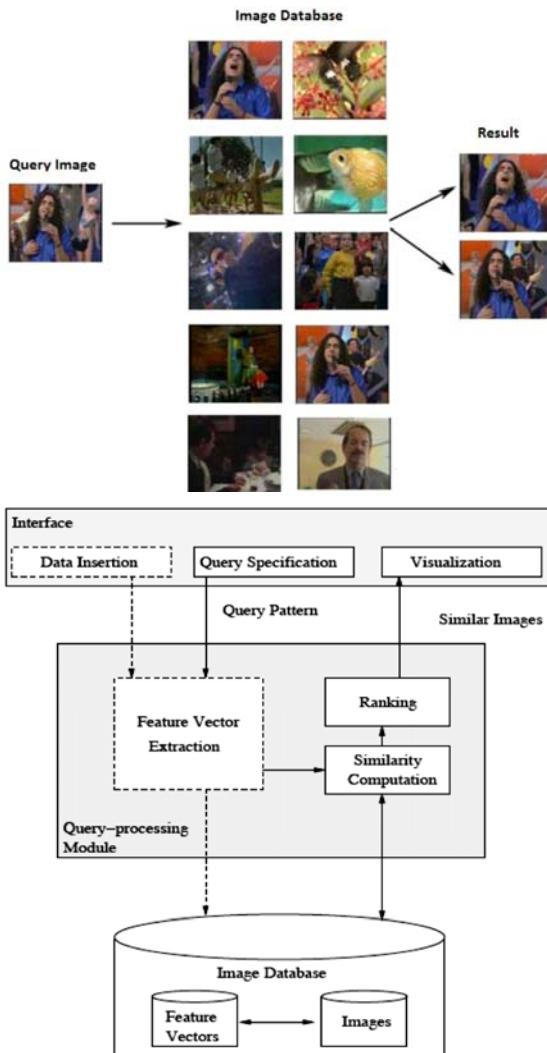
Advances in data storage and image acquisition technologies have enabled the creation of large image datasets. In order to deal with these data, it is necessary to develop appropriate information systems to efficiently manage these collections. Image searching is one of the most important services that need to be supported by such systems.

The shortcomings of textual metadata have been addressed by the so-called Content-Based Image Retrieval (CBIR) systems. Content-based image retrieval (CBIR) is a technique used for extracting similar images from an image database. Content-based image retrieval (CBIR), also known as query by image content (QBIC) and content-based visual information retrieval (CBVIR) is the application of computer vision to the image retrieval problem, that is, the problem of searching for digital images in large databases. The last decade has witnessed great interest in research on content-based image retrieval. This has paved the way for a large number of new techniques and systems and a growing interest in associated fields to support such systems.

Images are particularly complex to manage – besides the volume they occupy, retrieval is an application-and-context-dependent task. It requires the translation of high-level user perceptions into low-level image features (this is the so-called “semantic gap” problem).

Fig. 1 shows a typical architecture of a content-based image retrieval system. Two main functionalities are supported: data insertion and query processing. The data insertion subsystem is responsible for extracting appropriate features from images and storing them into the image database (see dashed modules and arrows). This process is usually performed off-line.

A typical CBIR solution requires the construction of an image descriptor, which is characterized by: (i) an extraction algorithm to encode image features into feature vectors; and (ii) a similarity measure to compare two images. Image descriptor can be one or combination of the following: (i) color descriptor; (ii) shape descriptor; (iii) texture descriptor. The similarity measure is a matching function, which gives the degree of similarity for a given pair of images as represented by their feature vectors, often defined as an inverse function of the distance (e.g., Euclidean), that is, the larger the distance value, the less similar the images are.



**Fig. 1.** A typical architecture of a content-based image retrieval system.

## 4. Advancements in CBIR

### 4.1. Early CBIR Techniques

“Content-based” means that the search will analyze the actual contents of the images. The term ‘content’ in this context might refer to colors, shapes, textures, or any other information that can be derived from the images itself. The very basic issue in designing a CBIR system is to select the most effective image features to represent image contents. Many low-level features include color features, such as color correlogram, color moments, color histogram and texture features. Color, texture and shape information have been the primitive image descriptors in content based image retrieval systems.

### 4.2. Advancements in CBIR

The image information is exploited jointly in image space, scale, and orientation domains can provide richer clues. This process involves two

phases. In the first phase, the face image is decomposed into different scale and orientation responses by convolving with multi-scale and multi multiorientation Gabor filters. In the second phase, LBP analysis is used to describe the neighboring relationship not only in image space but also in different scale and orientation responses. The LBP has been also used for texture segmentation, background modeling and detection, shape localization, interest region description, and biomedical image retrieval [12]. The versions of the LBP and the LDP in the open literature cannot adequately deal with the range of appearance variations that commonly occur in unconstrained natural images due to illumination, pose, facial expression, aging, partial occlusions, etc. [13-16]. In order to address this problem, the local ternary pattern (LTP) has been introduced for face recognition under different lighting conditions.

A novel image indexing and retrieval algorithm using local binary patterns (LBPs), local ternary patterns (LTPs) and local tetra patterns (LTrPs) for content-based image retrieval (CBIR) have increased the performance.

### 4.3. Local Binary Pattern (LBP)

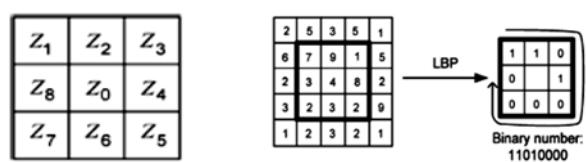
Local binary patterns (LBP) are a type of feature used for classification in computer vision. LBP is the particular case of the Texture Spectrum model proposed in 1990. LBP was first described in 1994. It has since been found to be a powerful feature for texture classification; it has further been determined that when LBP is combined with the Histogram of oriented gradients (HOG) classifier, it improves the detection performance considerably on some datasets.

The LBP operator was introduced by in for texture classification. Given a center pixel in the image, the LBP value is computed by comparing its gray value with its neighbors, as shown in Fig. 1, based on

$$LBP_{P,R} = p=1 P_2(p-1) f_1(g_p-g_c) \\ f_1(x) = 1, x \geq 0, \text{ otherwise}$$

where  $g_c$  is the gray value of the center pixel,  $g_p$  is the gray value of its neighbors,  $P$  is the number of neighbors, and  $R$  is the radius of the neighborhood.

The LBP feature vector, in its simplest form, is created in the following manner as shown in Fig. 2.



**Fig. 2.** Example of LBP feature vector creation.

The procedure is as follows. Divide the examined window into cells (e.g. 16x16 pixels for each cell). For each pixel in a cell, compare the pixel to each of its 8 neighbors (on its left-top, left-middle, left-bottom, right-top, etc.). Follow the pixels along a circle, i.e. clockwise or counter-clockwise. Where the center pixel's value is greater than the neighbor's value, write "1". Otherwise, write "0". This gives an 8-digit binary number (which is usually converted to decimal for convenience). Compute the histogram, over the cell, of the frequency of each "number" occurring (i.e., each combination of which pixels are smaller and which are greater than the center). Optionally normalize the histogram. Concatenate normalized histograms of all cells. This gives the feature vector for the window. The feature vector can now be processed using the support vector machine or some other machine-learning algorithm to classify images. Such classifiers can be used for face recognition or texture analysis.

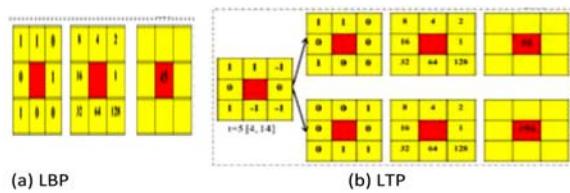
#### 4.4. Local Ternary Patterns (LTPs)

Local Ternary patterns are extensions of Local Binary Patterns.

Unlike LBP, it does not threshold the pixels into 0 and 1, rather it uses a threshold constant to threshold pixels into three values. LBP to a three-valued code called the LTP, in which gray values in the zone of width  $\pm t$  around  $g_c$  are quantized to zero, those above ( $g_c+t$ ) are quantized to 1, and those below ( $g_c-t$ ) are quantized to -1, i.e., indicator is replaced with three-valued function.

$$f_1(g_p, g_c, t) = \begin{cases} +1, & g_p \geq g_c + t \\ 0, & g_p - g_c < t \\ -1, & g_p \leq g_c - t \end{cases}$$

The binary LBP code is replaced by a ternary LTP code, as shown in Fig. 3.



**Fig. 3.** Calculation of the LBP and LTP operators.

In the LTP, the obtained ternary pattern is further coded into upper and lower binary patterns. The upper pattern is obtained by retaining 1 and replacing 0 for 1 and 0. Lower pattern is coded by replacing 1 with 1 and 0 for 1 and 0. In this way, each thresholded pixel has one of the three values. Neighboring pixels are combined after thresholding into a ternary pattern. Computing a histogram of these ternary values will result in a large range, so the ternary pattern is split into two binary patterns.

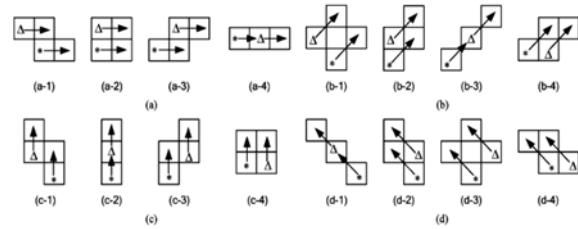
Histograms are concatenated to generate a descriptor double the size of LBP.

#### 4.5. Local Derivative Patterns (LDPs)

LDPs considered the LBP as the non-directional first order local pattern operator and extended it to higher orders (nth-order) called the LDP. The LDP contains more detailed discriminative features as compared with the LBP. To calculate the nth-order LDP, the (n-1)<sup>th</sup>-order derivatives are calculated along 00, 450, 900, and 1350 directions, denoted as  $I_{\alpha}^{(n-1)} (g_c) |_{\alpha=0,45,90,135}$ . Finally, n<sup>th</sup>-order LDP is calculated as

$$\begin{aligned} LDP_{\alpha}^n (g_c) &= p = 1P2 p - 1 f 2 I \propto n - 1 g c, \\ I &\propto n - 1 g p \mid P = 8 \\ f1(x) &= 1, \text{ if } x.y \leq 00, \text{ otherwise} \end{aligned}$$

The illustration of LDP templates and calculation is shown in Fig. 4.



**Fig. 4.** Illustration of LDP templates: (a) zero degree; (b) 45 degrees; (c) 90 degrees; (d) 135 degrees.

The detailed explanation along with figure, calculation of second-order LDP is shown in Fig. 5. The micro-patterns for second-order LDP is performed and shown only for zero

- a-8 neighborhoods of  $Z_0$
- b-.Template (a-1)Ref 1=  $Z_0$  bit=0
- c-.Template (a-2)Ref 1=  $Z_0$  bit=1
- d-.Template (a-3)Ref 1=  $Z_0$  bit=0
- e-.Template (a-4)Ref 1=  $Z_0$  bit=1
- f-.Template (a-1)Ref 2=  $Z_0$  bit=0
- g-.Template (a-2)Ref 2=  $Z_0$  bit=1
- h-.Template (a-3)Ref 2=  $Z_0$  bit=0
- i-.Template (a-4)Ref 2=  $Z_0$  bit=0

**Fig. 5.** Illustration of LDP templates for zero degree.

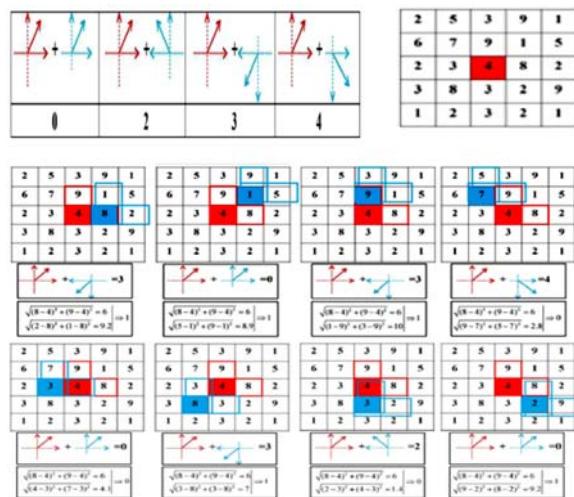
$LDP00 \ 2(Z_0) = 01010100$  is obtained by concatenating all the bits from above example. Similarly for  $45^\circ$ ,  $90^\circ$  and  $135^\circ$ , LDPs are given by  $LDP450 \ 2(Z_0) = 00101111$ ,  $LDP900 \ 2(Z_0) = 11010000$  and  $LDP1350 \ 2(Z_0) = 11000110$ . The second-order LDP for  $Z_0$  is  $LDP^2(Z_0) = 0101010000101111101000011000110$ .

#### 4.6. Local Tetra Pattern (LTrP)

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 LTrP 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. In addition, we propose a generic strategy to compute nth-order LTrP using  $(n-1)^{\text{th}}$ -order horizontal and vertical derivatives for efficient CBIR and analyze the effectiveness of our proposed algorithm by combining it with the Gabor transform.

The LBP, the LDP, and the LTP extract the information based on the distribution of edges, which are coded using only two directions (positive direction or negative direction). Thus, it is evident that the performance of these methods can be improved by differentiating the edges in more than two directions.

This observation has motivated us to propose the four direction code, referred to as local tetra patterns (LTrPs) for CBIR. A second-order LTrP that is calculated based on the direction of pixels using horizontal and vertical derivatives. It makes use of 00 and 900 derivatives of LDPs for further calculating the directionality of each pixel (Fig. 6).



**Fig. 6.** Calculating of tetra pattern bits for the center-pixel direction ‘1’ using the direction of neighbors. Tetra Pattern = 3 0 3 4 0 3 2 0 Pattern1 = 0 0 0 0 0 1 0; Pattern2=1 0 1 0 0 1 0 0; Pattern3 = 0 0 0 1 0 0 0 0  
Magnitude Pattern = 1 1 0 0 1 0 1.

For generating a tetra pattern, the bit is coded with the direction of neighbor when the direction of the center pixel and its neighbor are different, otherwise ‘0’. For the magnitude pattern, the bit is coded with ‘1’ when the magnitude of the center pixel is less than the magnitude of its neighbor, otherwise ‘0’.

#### 5. Research Challenges

CBIR is a challenging research problem till today. There are many novel methods in extracting features, describing images as a set of features,, matching and retrieval has been proposed and the process is still on. The implementation of CBIR systems raises several research challenges, such as:

- Formalisms need to be created to describe image content descriptions and related services. This formalism can guide the design and implementation of new applications based on image content.
- Not many techniques are available to deal with the semantic gap presented in image and their textual descriptions. New tools for marking/annotating images (and their regions) need to be developed. Better semantically enriched descriptions can be created by taking advantage of ontologies .
- Need for tools that automatically extract semantic features from images: extract high level concepts contained in multimedia data.
- Development of new data fusion algorithms to support text-based and content-based retrieval combining information of different heterogeneous formats .
- Finding new connections, and mining patterns. Text mining techniques might be combined with visual-based descriptions. New user interfaces for annotating, browsing and searching based on image content need to be investigated. Research in this area will require usability studies with practitioners.

#### 6. Conclusions

This paper has presented a brief overview of content-based image retrieval area. The CBIR is challenging research areas and there is a need for high performance. The basic principal of CBIR along with architecture is discussed. The novel approach referred as LTrPs for CBIR are discussed from literature. The LTrP encodes the images based on the direction of pixels that are calculated by horizontal and vertical derivatives. The performance improvement of the proposed method has been compared with the LBP, the LTP, and the LDP on grayscale images. In LTrP, only horizontal and vertical pixels have been used for derivative calculation. Results can be further improved by considering the diagonal pixels for derivative calculations in addition to horizontal and vertical directions. Due to the effectiveness of the proposed method, it can be also suitable for other pattern recognition applications such as face recognition, fingerprint recognition, etc.

The applications of CBIR are discussed along with potential research challenges in the area of CBIR are address. This may lead to new methods to

improve the performance of existing techniques and retrieval rate.

## Acknowledgements

This work was supported by the Science and Research Program of Nature Science Foundation of China under Grant (No. 60875006, 61162021), the Science and Technology Support Project of Gansu Province under Grant (No. 1304FKCA097) and Innovative Team Subsidize of Northwest University for Nationalities.

## References

- [1]. B. S. Manjunath and W. Y. Ma, Texture features for browsing and retrieval of image data, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 18, Issue 8, August 1996, pp. 837–842.
- [2]. M. Unser, Texture classification by wavelet packet signatures, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 15, Issue 11, November 1993, pp. 1186–1191.
- [3]. M. N. Do and M. Vetterli, Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance, *IEEE Transaction on Image Processing*, Vol. 11, Issue 2, February 2002, pp. 146–158.
- [4]. M. Kokare, P. K. Biswas, and B. N. Chatterji, Texture image retrieval using rotated wavelet filters, *Pattern Recognition Letters*, Vol. 28, Issue 10, July 2007, pp. 1240–1249.
- [5]. T. Ojala, M. Pietikainen, and D. Harwood, A comparative study of texture measures with classification based on feature distributions, *Pattern Recognition*, Vol. 29, Issue 1, January 1996, pp. 51–59.
- [6]. T. Ojala, M. Pietikainen, and T. Maenpaa, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 24, Issue 7, July 2002, pp. 971–987.
- [7]. Zou Hua Dong, and Hao Xiang Wang, A new automatic focusing algorithm and its application on vision measuring machine, *Applied Mechanics and Materials*, Vol. 397, 2013, pp. 1523–1528.
- [8]. Z. Guo, L. Zhang, and D. Zhang, Rotation invariant texture classification using LBP variance with global matching,” *Pattern Recognition*, Vol. 43, Issue 3, March 2010, pp. 706–719.
- [9]. S. Liao, M. W. K. Law, and A. C. S. Chung, Dominant local binary patterns for texture classification, *IEEE Transactions on Image Processing*, Vol. 18, Issue 5, May 2009, pp. 1107–1118.
- [10]. Z. Guo, L. Zhang, and D. Zhang, A completed modeling of local binary pattern operator for texture classification, *IEEE Transactions on Image Processing*, Vol. 19, Issue 6, June 2010, pp. 1657–1663.
- [11]. H. Lategahn, S. Gross, T. Stehle, and T. Aach, Texture classification by modeling joint distributions of local patterns with Gaussian mixtures, *IEEE Transactions on Image Processing*, Vol. 19, Issue 6, June 2010, pp. 1548–1557.
- [12]. G. Zhao and M. Pietikainen, Dynamic texture recognition using local binary patterns with an application to facial expressions, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 29, Issue 6, June 2007, pp. 915–928.
- [13]. Wang Haoxiang, Ferdinand Shkjezi, and Ela Hoxha, Distance metric learning for multi-camera people matching, in *Proceedings of the IEEE Sixth International Conference on Advanced Computational Intelligence (ICACI)*, 2013, pp. 140–143.
- [14]. Z. Lei, S. Liao, M. Pietikäinen, and S. Z. Li, Face recognition by exploring information jointly in space, scale and orientation, *IEEE Transactions on Image Processing*, Vol. 20, Issue 1, January 2011, pp. 247–256.
- [15]. X. Tan and B. Triggs, Enhanced local texture feature sets for face recognition under difficult lighting conditions, *IEEE Transactions on Image Processing*, Vol. 19, Issue 6, June 2010, pp. 1635–1650.
- [16]. G. S. Chen, H. Y. Tang, D. Q. Zhang, Y. X. Jiao, & H. X. Wang, Existence of three solutions for a nonlocal elliptic system of-Kirchhoff type, *Boundary Value Problems*, Issue 1, 2013, pp. 1–9.
- [17]. M. Subrahmanyam, R. Balasubramanian and R. P. Maheshwari, Local tetra patterns: a new feature descriptor for content-based image retrieval, *IEEE Transactions on Image Processing*, Vol. 21, Issue 5, May 2012, pp. 2874–2886.