Available Online at www.ijcsmc.com

International Journal of Computer Science and Mobile Computing

A Monthly Journal of Computer Science and Information Technology

ISSN 2320–088X

IJCSMC, Vol. 3, Issue. 12, December 2014, pg.466 – 473

RESEARCH ARTICLE

Image Retrieval Based on Quad Chain Code and Standard Deviation

Prof. Dr. Tawfiq A. Al-assadi [1], Prof. Dr. Israa Hadi Ali [2], Abdul kadhem Abdul Kareem Abdul kadhem [3]

[1](Information Technology College/ Babylon University, Iraq)
[2](Information Technology College/ Babylon University, Iraq)
[3](Software Department, Information Technology College/ Babylon University, Iraq)
[1] tawfiqasadi63@yahoo.com; [2] israa_hadi1968@yahoo.com; [3] kazum2006k@yahoo.com

Abstract— In this paper we present a new idea for image retrieval depending on Quad chain code and standard deviation. The proposed system consist of two stages ,the first one is constructing a database of image training by dividing each image into number of blocks with size 8*8 pixel based on Quad chain code and the structure of database for each block of image consisting of one record with two columns (first column contain a set of chain code and the second column contain standard deviation of block pixels).In the second stage the test image is divided into a number of block based on quad chain code and find standard deviation for each block and apply a matching operation for each image in database to find what is the most matching image.

Keywords— feature extraction, chain code, image retrieval, standard deviation

I. INTRODUCTION

Image Retrieval System offers an effective and sophisticated way for maintaining and managing image databases. Image retrieval (IR) is one of the most exciting and fastest growing research areas in the field of medical imaging, aerial/satellite image analysis and retrieval, etc. The goal of Content-Based Image Retrieval (CBIR) is to retrieve images from a database that are more similar to an image seen as an image test (query). In CBIR, for each image in the training set image and test image, features are extracted and compared to the features of the test image. A CBIR method typically converts an image into a feature vector representation and matches with the images in the database (training set) to know most similar images [1]. Image retrieval problem, such as recognizing, classification the similar images, the need for databases, the semantic gap, and retrieving the desired images from huge number of images are the keys to improve. In some cases, can be considered as a pattern classification problem, where each image is supposed as ground truth returns to a specific class. Then query by test image is to find the class and return images within that class. But usually there is no class label available so this can only be achieved by comparing among images by using the similarity measure [2, 3]. A feature that is extracted from images is the basis of content-based image retrieval. In general, features may include visual features (texture, color, shape, faces). However, within the visual feature scope, the features can be classified as general features and domain specific features. The first one include texture, color, and shape features while the second one is application dependent and may include, for example, human faces and finger prints. The domain-specific features are better covered in pattern recognition literature and may include much domain knowledge [4]. In this paper the proposed system must first build database as follows: Finding blocks

with size 8*8 depend on Quad chain code, compute standard deviation for each block, and index the images in the database using a vector for each image that contains (stream of chain code for each block, the standard deviation for each block). The next step is to check the test image (query) with the database and apply similarity measures for matching operations to find if there are one or more images in the database that is more matching with the test image.

1.1. Chain code.

Chain code is used to represent the border shapes of the image through a connected sequence of straight-line segments of specified length and direction, this representation is based on 4- or 8-connectivity of the segments. The 4-connectivity allows only the horizontal and vertical movements between neighbour pixels, whereas the 8-connectivity uses the diagonal movement [5], as shown in fig (1).

In the traditional freeman chain code the direction of each segment is coded by using a numbering system. A boundary code formed as a sequence of directional numbers is referred as a Freeman Chain Code. The chain code of the shape's border depends on the start point. Working with code numbers offers a unified tool to analyse the shape of the boundary. The problem of 4-connectivity is that we lose the diagonal points where these points are useful in most of the image applications. So, in order to skip this problem of 4-connectivity here we use 8-connectivity. In 8- connectivity each code can be considered as the angular directions, multiplied by 45 degree that we must move to go from one contour pixel to the next [6]. Showing example in fig (2)

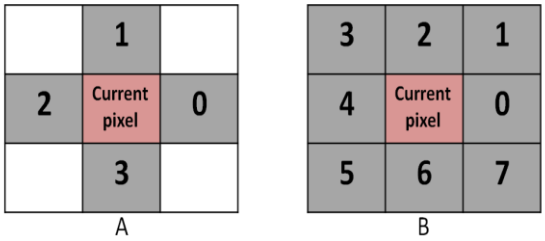
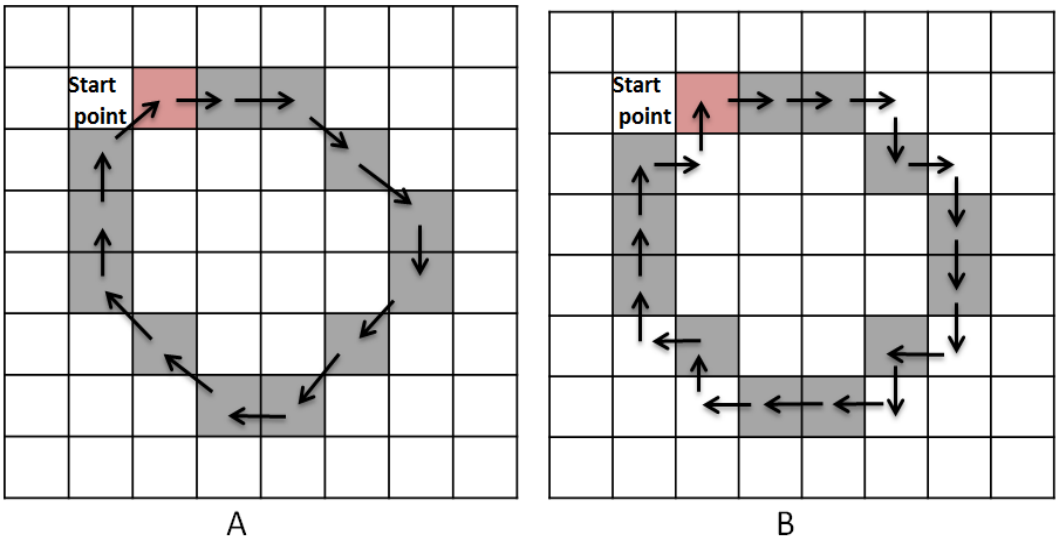


Fig. 1 A. 4- connectivity chain code . B. 8- connectivity chain code.



A: The chain code 8- connectivity = 0,0,7,7,6,5,5,4,3,3,1,3,1
B:the chain code 4- connectivity=0,0,0,3,0,3,3,3,2,3,2,2,2,1,2,1,1,1,0,1

Fig. 2 Example of extract chain code (clockwise) .

in this paper we will develop the traditional freeman chain code with 8-connectivity to quad chain code, and find the chain code for an image depending on one of the measures which is the similarity between two adjacent vectors (quad pixels).

1.2. Feature extraction.

The feature is a function of one or more measurements, each of which specifies some measurable properties of an object, and is computed such that it quantifies some significant characteristics of the object. For each image in the image database, its features are extracted as a vector that is saved in the feature database. When a test image (query) comes in, its feature space will be compared with those in the feature database one by one

and the similar images with the smallest feature distance will be retrieved [7]. Feature extraction includes the image features to a distinguishable extent. Average RGB, Color Moments, Co-occurrence, Local Color Histogram, Global Color Histogram and Geometric Moments are used to extract features from the test image. Feature matching, on the other hand, includes matching the extracted features to get results that show visual similarities. Feature vectors are calculated for the given image. The Euclidean distance is used as the default implementation compared between two vectors feature. If the distance between vector feature of the query image and database of images is small enough, the corresponding image in the database is considered as a match to the test image. The colour values of the image are one of the most important features of the images. Color features are defined a subject to a specific color space or model. The common moments mean, standard deviation and skewness [8]. In this paper we will extract the features (mean, standard deviation) for each block of images and use these features for matching the operation between image query and database, the mean and standard deviation formula can be defined as follows.

$$MEAN = \mu = \frac{1}{N}\sum_{k=1}^{N} pixel(k)$$

$$STANDARD\ DEVIATION = \sigma = \sqrt{\frac{1}{N}\sum_{k=1}^{N}(pixel\ (k) - \mu\ )^2}$$

II. PROPOSED SYSTEM

The proposed system consist of two stages, the first one as illustrated in Fig (3) is build the database from large number of images and extract features from the image depending on (QCC) Quad Chain Code standard deviation and save it as a vector for each image in the database. The second part as illustrated in Fig (4) is checking the image query with the database to retrieval image that matches the most.

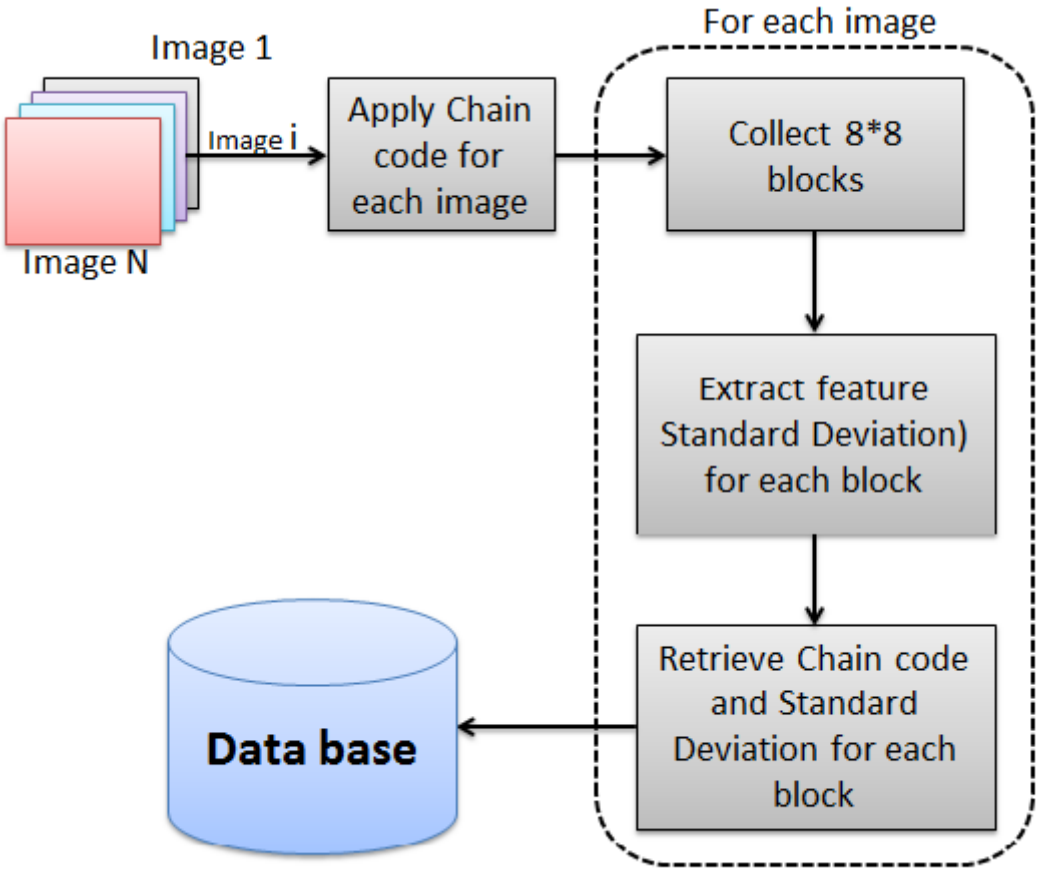


Fig (3) block diagram for build the database.

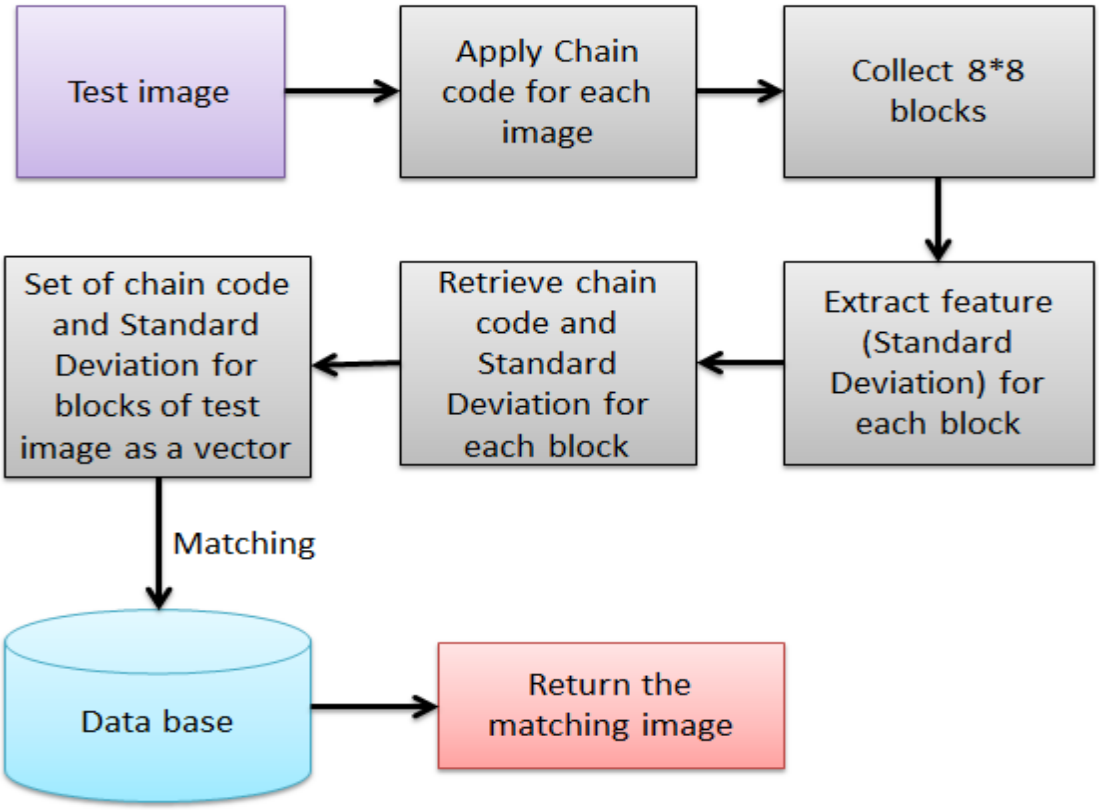


Fig (4) block diagram of image retrieval.

2.1. The first stage (build the database).
In this stage there are several steps to build database by extracting the feature (standard deviation) from images, it consist of:

2.1.1. Read Images.
In this step will be read number of images that used to build database.

2.1.2. Apply Chain code (quad).
In this paper we are using the 8-connected FCC (freeman chain code) each code can be considered as the angular direction in multiple of 45 degree that we must move to go from on contour pixel to the next. The Proposed method is that we will develop the chain code by using Quadruple pixel (vector of 4 pixels), and find the chain code for each image depending on the measure of similarity between two adjacent vectors. Figure (5) show an example that illustrates how the Quad chain code computed in an image.

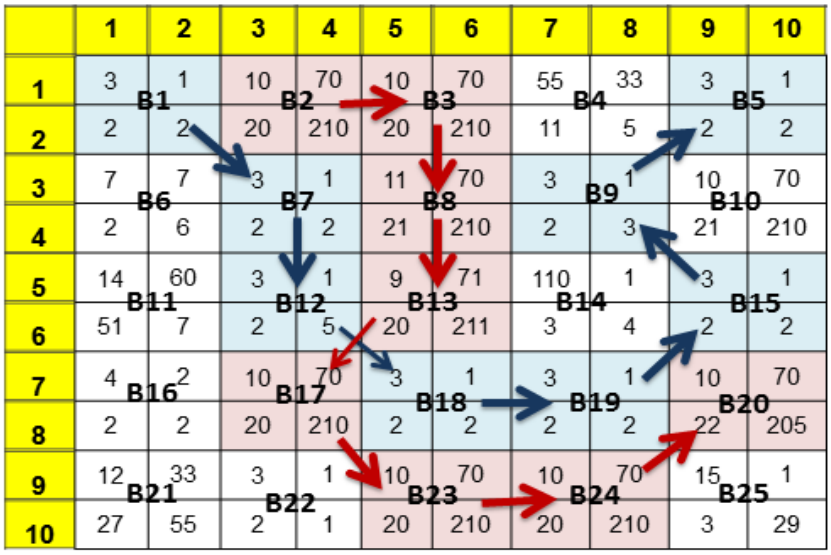
2.1.3. Collect 8*8 blocks.
In this step we will be Collecting the quad pixels that were extracted from the above step for equip block with size 8×8 .each quad pixel of the 8×8 block (16 Quad pixels,64 pixels) is similar with each other. The similarity measure that can be applied in the proposed system that is used to find alike values between Quad pixels or blocks Quartet is Cosine Similarity, as shown in the following equation and example:
If Block1 and Block2 are two Quad pixels (vector), then

$$\text{Cos}(\text{Block1}, \text{Block2}) = (\text{Block1} \bullet \text{Block2}) / \|\text{Block1}\| \, \|\text{Block2}\|$$

Where • indicates the vector dot product and || Block || is the length of the vector Block.
For example let Block1=(3,1,2,2) and Block2=(3,1,2,5) then the cosine similarity between Block1 and Block2 is:

$$\text{Block1} \bullet \text{Block2} = 3*3 + 1*1 + 2*2 + 2*5 = 24$$
$$\|\text{Block1}\| = (3*3+1*1+2*2+2*2)0.5 = (18)\ 0.5 = 4.242$$
$$\|\text{Block2}\| = (3*3+1*1+2*2+5*5)\ 0.5 = (39)\ 0.5 = 6.244$$
$$\text{Cos}(\text{Block1},\text{Block2}) = 0.90610.$$

As long as the cosine similarity value is closer to 1 than that mean the block is more alike and Vice versa.



The first quad chain code in location (1,1)=7,6,7,0,1,3,2
The blocks = B1,B7,B12,B18,B19,B15,B9,B5
The second quad chain code in location (1,2)=0,6,6,5,7,0,1
The blocks= B2,B3,B8,B13,B17,B23,B24,B20

Fig (5). Quad chain code example.

2.1.4. Extract standard deviation.
In this step we find the standard deviation for each block of the image. Standard deviation represents contrast of image. Standard deviation is calculated for pixels in each block of an image to get feature vector which is used for image retrieval [9].

2.1.5. Retrieve chain code and Standard Deviation for each block.
In this step we will store the stream of chain code and the value of standard deviation for each block of image in record with two columns. The structure of the block is shown in the following example.

Chain code for block	Standard deviation
7,0,0,5,0,0,3,2,2,2,1,1,3,3,4,7	57

After this operation, this record will be send to database and collect another block of quad pixels that is more similar by going back to step 2.1.3 until it covers all blocks in the image.

2.1.6. Database.
In this step we collect the records of the image (chain code, standard deviation) that was extracted from the above step into a vector, that mean each image will be converted to a vector, as the following table(1) explain the database.

TABLE (1). Structure of database.

images	Block 1		Block 2		……	Block N	
	Chain code	STD	Chain code	STD	…….	Chain code	STD
Image1	0077523404235100	75	0001002377564446	215	…….	0001234757471002	122
Image2	7722100470017771	13	3102657772510375	26	…….	2147777111000454	17
Image3	3157111161000211	128	1377544444301673	79	…….	6657663667752001	244
. . . .	. . . .	. . . .	. . . .	. . . .	…….	. . . .	. . . .
ImageN	1112700117700001	212	5564700012776661	117	…….	7145552000147632	25

2.2. The second stage (Image Retrieval).
In this stage there are several steps to test image with database to retrieval images that are more matching, the second stage consist of :

2.2.1. Test image.
In this step we will input the query image for testing with database.

2.2.2. Apply chain code for test image.
In this step we apply chain code for test image to find the data (quad pixels) that is more similar, such as done the example in fig (5).

2.2.3. Collect 8*8 blocks.
In this step we equip 64 pixels that are similar to each other to a block such as in step 2.1.3 above.

2.2.4. Extract standard deviation.
After collecting each 64 pixels to a block we find the standard deviation for blocks of the image.

2.2.5. . Retrieve chain code and Standard Deviation for each block.
In this step we will store the stream of chain code and the value of standard deviation for each block of test image in record with two columns such as done in step 2.1.5 above.

2.2.6. Set of chain code and Standard Deviation for blocks of test image as a vector.
In this step we convert the test image to a one vector, the vector contains number of records each record contains two fields, the first one contains the stream of chain code and the second contains the standard deviation.

2.2.7. Image retrieval.
In this step we check the vector of test image with vectors of database by using the similarity measure to find number of blocks in database that is more alike with blocks of test image .the matching operation first checked the similarity of stream chain code and then checked the similarity between standard deviation for blocks. The final result returns images from database that contains large number of matching blocks.

III. ALGORITHMS AND EXPERIMENTAL RESULT

3.1. Algorithms

1. Build database algorithm

- **Input: number of image training (N).**
- **Output: numbers of vectors.**

Step0: **begin**.
Step1: index=1
Step2: **Repeat**
Step3: **Read** image [index] then convert it to matrix.
Step4: d=1
Step5: **if** found 8*8 pixels (block) with more similarity **then**
Step6: **begin**.
Step7: find STD for block.
Step8: set stream of chain code.
Step9: save chain code and STD in record.
Step10: save record in vector [index] [d]
Step11: d=d+1
Step12: **go to** step 5
Step13: **end if**
Step14: index =index+1
Step15: **until** index = N.
Step 16: **end**

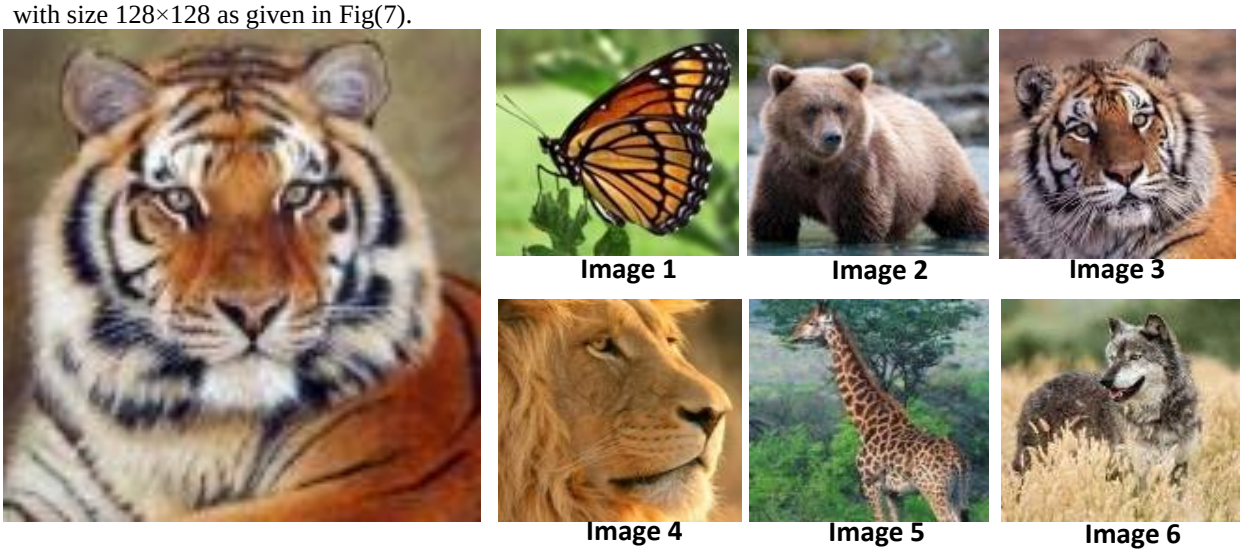
2. Image retrieval algorithm

- **Input: test image.**
- **Output: image retrieval.**

Step0: **begin**.
Step1: **Read** test image then convert it to matrix.
Step2: d=1
Step3: **if** found 8*8 pixels (block) with more similarity **then**
Step4: **begin**.
Step5: find STD for block.
Step6: set stream of chain code.
Step7: save chain code and STD in record.
Step8: save record in vector [k] [d]
Step9: d=d+1
Step10: **go to** step 5
Step11: **end if**
Step12: index=1
Step 13: **Repeat**
Step14: save number of block matching between vector [k] and vector [index] in matching table [index].
Step15: index =index +1
Step16: **until** index = N.
Step 17: return image that largest in matching table.
Step 16: **end**

3.2. Experimental result.

Experiments of proposed method in this paper are performed on two training image sets, the first database contains 6 color images with size 128 ×128 as given in Fig.(6) .and the other database contains 6 color image with size 128×128 as given in Fig(7).



a. Test image (Query) b. Database image

Fig (6). The first database

the image database is used to retrieve the relevant images based on test image(query) and retrieve the image that contain the large number of matching blocks in database. The results explained in table (2) is to retrieve image 3 because it have large number of matching block when number of blocks for test image is (104) block and the similarity ratio for collected blocks from image database is 99% and the similarity ratio between blocks of test image and blocks of image database is 90%.

TABLE (2) . The Experimental rustle for database1

images	No. of Block for each image	No. of matching block
Image1	107	80
Image2	136	79
Image3	107	98
Image4	172	89
Image5	74	61
Image6	140	76

a. Test image (query). b. Database image.

Fig (7). The second database

After applying the proposed system for Database 2 The results explained in table (3) is to retrieve image 4 because it have large number of matching block when number of blocks for test image is (197) block and the similarity ratio for collected blocks from image database is 99% and the similarity ratio between blocks of test image and blocks of image database is 90%.

TABLE (3) . The Experimental rustle for database2

images	No. of Block for each image	No. of matching block
Image1	**122**	**52**
Image2	**77**	**47**
Image3	**93**	**46**
Image4	**130**	**94**
Image5	**131**	**57**
Image6	**171**	**56**

IV. CONCLUSIONS

The huge increase in sizes of image database has helped to spur the development of effective and efficient retrieval systems. In this paper we present an algorithm for retrieving the image from database based on quad chain code and standard deviation. The proposed system to image retrieval depend on two things, the first one is the locations of pixels which corresponding to stream chain code, the second one is the colors values which corresponding to standard deviation. The quad chain code used to choose some regions of an image that is similar with each other to collect number of blocks of image, each block with size 64 pixels. The standard deviation tells us something about the contrast of blocks of image. The proposed system extracts the standard deviation for blocks of test image and compares them to those of database images by using matching and comparison algorithms. The matching process for each of the stream chain code and standard deviation is to find images from database that contains large number of matching blocks for retrieving.

REFERENCES

[1] Vadhri Suryanarayana, Dr. M.V.L.N. Raja Rao, Dr. P. Bhaskara Reddy,Dr. G. Ravindra Babu, "IMAGE RETRIEVAL SYSTEM USING HYBRID FEATURE EXTRACTION TECHNIQUE", *International Journal of Computer Science & Information Technology (IJCSIT) Vol 4, No 1, Feb 2012*.

[2] Xiang Sean Zhou, Ira Cohen, Qi Tian, Thomas S. Huang, "Feature Extraction and Selection for Image Retrieval " *Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana Champaign, Urbana, IL 61801*.

[3] Nishant Singh, Shiv Ram Dubey, Pushkar Dixit, Jay Prakash Gupta," SEMANTIC IMAGE RETRIEVAL USING MULTIPLE FEATURES ", *Computer Science & Information Technology (CS & IT), 06, pp. 277–284, 2012.*

[4] Yong Rui and Thomas S. Huang," Image Retrieval: Current Techniques, Promising Directions, and Open Issues", *Journal of Visual Communication and Image Representation 10, 39–62 , Article ID jvci.1999.0413 (1999).*

[5] Tawfiq A. Al-Asadi, Fanar Ali Joda, "Removing Spatial Redundancy from Image by Using Variable Vertex Chain Code", *EUROPEAN ACADEMIC RESEARCH Vol. II, Issue 1/ April 2014.*

[6] Pulipati Annapurna, Sriraman Kothuri, Srikanth Lukka, " Digit Recognition Using Freeman Chain Code", *International Journal of Application or Innovation in Engineering & Management (IJAIEM), Volume 2, Issue 8, August 2013.*

[7] Aman Chadha, Sushmit Mallik, Ravdeep Johar, "Comparative Study and Optimization of Feature-Extraction Techniques for Content based Image Retrieval", *International Journal of Computer Applications (0975 – 8887)Volume 52– No.20, August 2012.*

[8] Dong ping Tian, "A Review on Image Feature Extraction and Representation Techniques", *International Journal of Multimedia and Ubiquitous Engineering ,Vol. 8, No. 4, July, 2013.*

[9] fazal-e-malik ," mean and standard deviation features of color histogram using laplacian filter for content-based image retrieval", *Journal of Theoretical and Applied Information Technology,15th December 2011. Vol. 34 No.1*

