International Journal of Computer Applications (0975 – 8887)
Volume 141 – No.7, May 2016

Content based Image Retrieval using Color and Texture

Tasneem Mirza
Thadomal Shahani Engineering College, Affiliated to Mumbai University, India

Rishabh Indoria
Thadomal Shahani Engineering College, Affiliated to Mumbai University, India.

Geet Dalwani
Thadomal Shahani Engineering College, Affiliated to Mumbai University, India

Krish Jain
Thadomal Shahani Engineering College, Affiliated to Mumbai University, India.

ABSTRACT
A Content Based Image Retrieval System is a computer system for browsing, searching and retrieving images from a large database of digital images .Most common methods of image retrieval utilize some method of adding meta data such as captioning, keywords or description to the images so that retrieval can be performed over the annotation words. Content Based Image Retrieval (CBIR) deals with retrieval of images based on visual features such as color, texture and shape. This paper presents retrieval of images based on color and texture using various proposed algorithms.

General Terms
Image Retrieval

Keywords
Metadata, Content Based Image Retrieval

1. INTRODUCTION
Initially Text Based Image Retrieval (TBIR) systems were only used where the search is based on annotation of images. It searches for similar text surrounding the image as given in the query string. However, it is sometimes not easy to express the visual content of images in words and TBIR ends up in producing irrelevant results.

Nowadays, along with TBIR, CBIR is also used which is more accurate and efficient.

This paper presents Content Based Image Retrieval (CBIR), also known as Query by Image Content (QBIC) means that search analyzes the contents of the image rather than the metadata such as tags, keywords or descriptions associated with the image. The term “content” in this context refers to color and texture [1]. The objective of Content Based Image Retrieval (CBIR) is to efficiently retrieve images that satisfy a user’s criteria of similarity by exploiting multiple features addressing different image properties. The use of multiple features confronts the user with multiple difficulties. First, a deeper understanding of the feature functions which are implemented is required. Second, the user needs to understand how these feature functions are to be combined and further specification is to be provided by assigning thresholds and weights.

2. SYSTEM DESCRIPTION

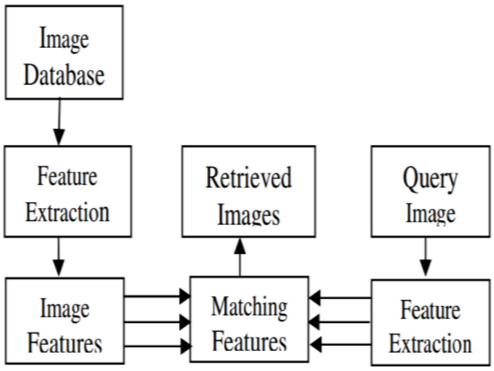

Fig 1: Block Diagram of CBIR

Block Diagram of CBIR shows the typical architecture of Content Based Image Retrieval System (CBIR). The query processing module extracts feature vectors from the query image and database images and applies a metric such as Euclidean distance to evaluate the similarity between query image and database images by comparing similar features of images [2]. Next, it ranks the database images in decreasing order of similarity to query image before it is displayed to the user.

3. COLOR FEATURE EXTRACTION
Color is a basic visual attribute for human perception and computer vision [3] and is one of the most widely used feature in image retrieval. Color Feature Extraction is based on selected color space. In the paper RGB and HSV color space have been used.

3.1 RGB Color Model
The RGB color model is a summative color model in which red, green and blue light are combined together in various ways to reproduce an array of colors. The model is summative in the sense that three light beams are summed or added together to make the final spectrum [4][5]. The RGB model is mainly used in sensing, representation and display of images in electronic system such as television or computer.

3.2 HSV Color Model
It is one of the most common cylindrical representations in an RGB color model. The Hue (H) represents the dominant spectral component in its pure form, as in green, red or yellow. Adding white to the pure color changes the color, the less white, the more saturated the color is. This corresponds to Saturation (S). The value (V) corresponds to the brightness of the color.

The algorithm for color feature extraction is discussed below.

3.3 Color Feature Extraction Using Mean

Steps to be followed to extract the color feature.

1. Read query image from user.
2. Extract the RGB components from the image and calculate their means individually.
3. The RGB component of each pixel is compared with the results obtained in step 2 for each component.
4. If the comparison performed in step 3 results true, then RGB component of each pixel is assigned to RH,GH and BH respectively otherwise, to RL,BL and GL respectively.
5. Calculate the mean of RH, RL, GH, GL, BH, BL and store it into a 256 bit color feature vector in database.
6. Calculate the similarity measurement of query image with the image present in database using Euclidean distance.
7. Retrieve the images based on a minimum distance.

4 TEXTURE FEATURE EXTRACTION

An image texture gives us information about the spatial arrangement of color or intensities in an image or selected region of an image. The algorithm for texture feature extraction is discussed below.

4.1 Texture Feature Extraction Using Standard Deviation

Standard Deviation is a measure that is used to quantify the variation or dispersion of a set of data values [6]. It is given as:

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2} \qquad (1)$$

Where variance, $\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$, N is the number of images in the database, μ is the standard deviation, x_i denotes the individual values.

Steps to be followed to extract the texture feature.

1. Read the query image from the user.
2. Extract the RGB component from the image and extract their means individually.
3. The RGB component of each pixel is compared with the results obtained in step 2 for each component.
4. If the comparison performed in step 3 results true, then RGB component of each pixel is assigned to RH, GH and BH respectively otherwise, to RL, GL and BL respectively.
5. Calculate the mean of RH, RL, GH, GL, BH, and BL and store it into a single feature vector w [0...5].
6. Repeat step 3 for all images in database.
7. Calculate the standard deviation for RH, RL, GH, GL, BH, and BL and store it into 3 bit texture feature vector.
8. Compare similarity matching of query image and database image using distance metrics.
9. Retrieve the images based on a minimum distance.

5 COMBINING COLOR AND TEXTURE FEATURES

The retrieval result using only single feature may be inefficient. It may either retrieve images not similar to query image or may fail to retrieve images similar to query image. Hence, to produce efficient results, combinations of color and texture features have been used in the paper. The similarity between query and target image is measured from two types of characteristic features which includes color and texture features as follows:

The 256 bit color feature vector obtained using the color algorithm discussed above is appended to the 3 bit texture feature vector using the texture algorithm discussed above. Hence a 259 bit feature vector which is computed for each image in the database .Using similarity measurement the combined system finds the similar image matching the query image from the database.

6 SIMILARITY MEASUREMENT

The Euclidean distance is used for similarity measurement. It allows the feature set to be in normalized form. It is given by:

$$Euclidean\ Distance = \sqrt{\sum_{i=1}^{n}[Q_i - D_i]^2}$$
(2)

Where 'n' is the number of images in the database, Q_i denotes i^{th} query image and D_i denotes the image selected to be compared from the database.

7 PERFORMANCE EVALUATION

The performance of the image retrieval of the CBIR system can be measured in terms of its recall and precision. Recall measures the ability of the system to retrieve all the models that are relevant, while precision measures the ability of the system to retrieve only the models that are relevant.

$$Precision = \frac{Number\ of\ relevant\ images\ retrieved}{Total\ number\ of\ images\ retrieved} \qquad (3)$$

$$Recall = \frac{Number\ of\ relevant\ images\ retrieved}{Total\ number\ of\ relevant\ images} \qquad (4)$$

8 RELEVANCE FEEDBACK

Relevance Feedback was introduced in CBIR to improve the performance by human intervention. The CBIR System is based on similarity rather than on exact match, and the results are given to the user. The user gives feedback or judgments on the retrieved results. It is based on two value assessment, relevant or not relevant. If the user feedback is relevant, then the loop stops otherwise, it continues till user gets satisfied with the results [7].

9 EXPERIMENTAL SETUP

The discussed image retrieval methods are implemented using Java JDK 1.8 on Intel Core I5 Processor with 4GB of RAM.

Data Set: WANG [8] database of 100X100 images consisting of 100 categories has been used. Each category contains 100 images.

Feature Set: It comprises of color and texture descriptors for each pixel of an image.

10 EXPERIMENTAL RESULTS AND CONCLUSION

Different Color and Texture algorithms proposed in the paper have been used to develop a framework to implement Content Based Image Retrieval (CBIR) System with a database of 10,000 images. There are 100 different categories containing 100 different images.

Query image is given to CBIR System by selecting an image from the 10,000 images in the GUI and found the similar image using Color and Texture descriptors.

First, the CBIR System using Color using Mean has been implemented. The implementation of the following is shown in the GUI depicted below:

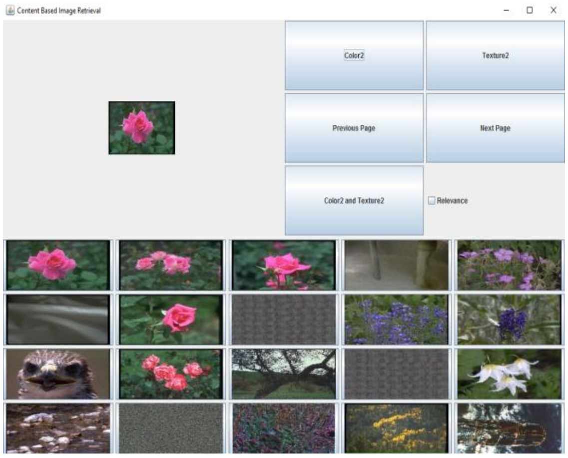

Fig 2: CBIR using Color using Mean

Second, the CBIR System using Texture using Standard Deviation has been implemented. The implementation of the following is shown in the GUI depicted below:

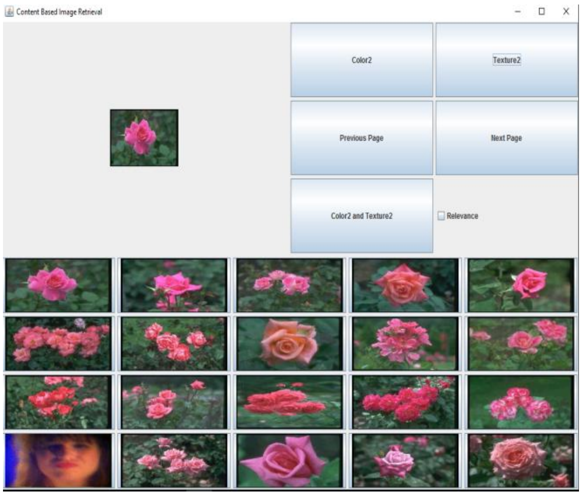

Fig 3: CBIR using Texture using Standard Deviation

Finally, the CBIR System using both Color and Texture using mean and standard deviation respectively has been implemented.

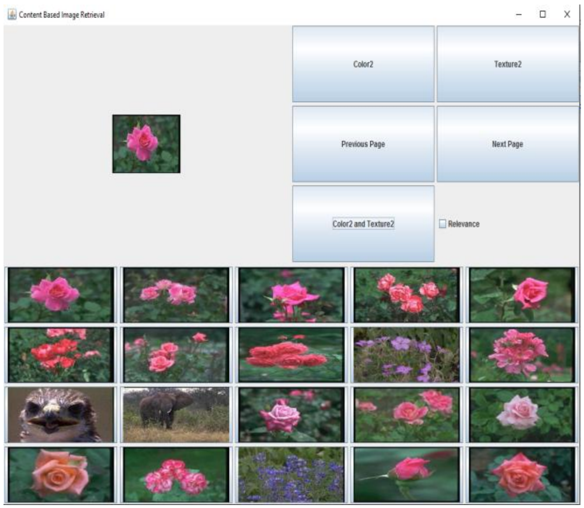

Fig 4: CBIR using Color and Texture using Mean and Standard Deviation

The average precision and recall of the proposed algorithms presented in the paper is represented in the tables given below for some of the categories from dataset.

Table 1. The average precision for all methods

Category	Proposed Color Algorithm	Proposed Texture Algorithm	Combination of Color and Texture
Butterflies	0.40	0.45	0.55
Animals	0.60	0.50	0.80
Flowers	0.60	0.95	0.90
Fruits	0.40	0.40	0.45
Nature	0.65	0.70	0.75
Doctors	0.80	0.80	0.85
Earth	0.55	0.50	0.65
Text	0.70	0.95	1.00
Railway Track	0.60	0.50	0.80
Pattern	0.70	0.95	0.95
Average Precision	0.6	0.67	0.77

Table 2. The average recall for all methods

Category	Proposed Color Algorithm	Proposed Texture Algorithm	Combination of Color and Texture
Butterflies	0.08	0.09	0.11
Animals	0.12	0.10	0.16
Flowers	0.12	0.19	0.18
Fruits	0.08	0.08	0.09
Nature	0.13	0.14	0.15
Doctors	0.16	0.16	0.17
Earth	0.11	0.10	0.13
Text	0.14	0.19	0.20
Railway Track	0.12	0.10	0.16
Pattern	0.14	0.19	0.19
Average Recall	0.108	0.134	0.154

11 ACKNOWLEDGEMENTS
Thanks and gratitude towards the guide and all the people, who have helped directly and indirectly.

12 REFERENCES
[1] Eakins, John; Graham,Margaret. "Content Based Image Retrieval" Northumbria at Newcastle. Retrieved 2014-03-10.

[2] J R Smith, "Integrated spatial and feature image system:Retrieval analysis and compression :[Ph D dissertation], Columbia University, New York, 1997.

[3] Rui Y, Huang T S, Chang S F. Image Retrieval: current techniques, promising directions and open issues, Journal of Visual Communication and Image Representation, 1999, 10(I): 39-62.

[4] Charles A. Poynton (2003). Digital Video and HDTV: Algorithms and Interfaces. Morgan Kaufmann. ISBN 1-55860-792-7.

[5] Nicholas Boughen (2003). Lightwave 3d 7.5 Lighting. Wordware Publishing, Inc.ISBN 1-55622-354-4.

[6] Bland, J.M.; Altman, D.G. (1996). "Statistics notes: measurementerror". BMJ **312**(7047):1654. doi:10.1136/bmj.312.7047.1654. PMC 2351401. PMID 8664723.

[7] Pushpa B. Patil, Manesh B. Kokare, "Relevance Feedback in Content Based Image Retrieval: A Review", Journal of Applied Computer Science & Mathematics, no. 10(5) /2011, Suceava.

[8] P.S. Hiremath, JagdeeshPujari. "Content Based Image Retrieval based on Color, Texture and Shape features using Image and its complement" International Journal of Computer Science and Security, Volume (1): Issue (4).