Unsupervised Content Based Image Retrieval at Different Precision Level by Combining Multiple Features

S. M. Zakariya¹ and Mohd Atif Jamil²
¹Electrical Engineering Section, ²Mechanical Engineering Section, University Polytechnic, Faculty of Engineering & Technology, Aligarh Muslim University, India
E-mail: s.zakariya@gmail.com

Abstract. Image retrieval is a procedure of finding appropriate images in the image database. There are two types of image retrieval systems in common practice. These are the text-based image retrieval (TBIR) system and content-based image retrieval (CBIR) system. The content based system is proven to be more effective in which the visual contents of the images are extracted and described by multi-dimensional feature vectors. In this work, several models are developed by combining different image features in a combination of two and three. To begin with, three different models based on the combination of two features, viz., color with shape, shape with texture, and color with texture are designed. A three features based model is considered with color, shape, and texture in the next step. The retrieval rate of the mentioned models is assessed in terms of precisions. The results are obtained using COREL standard database. This study shows that the images can be better retrieved using three features based model in contrast to models using two features.

1. Introduction

The users can access the data in a variety of media forms on the internet. The millions of thousands of digital images have been estimated on the World Wide Web. This produces a requisite for the development of innovative systems for healthier storage and efficient retrieval of images [12]. Content-based image retrieval (CBIR) is any growth that on a central level supports to constitute the databases of the digital image by using their visual substance. According to this explanation, whatever fluctuating from a strong image description system to an image resemblance task comes in the understanding of CBIR [17]. This representation of CBIR as an area of study places it at a special connection inside the intellectual group. While the strong point in resolving the vital problem of in a good physical shape image understanding, discern the researcher from distinct credentials, such as information retrieval, information theory, computer vision, human-computer interaction, machine learning, Web and data mining, database systems, and statistics assisting and attaching with the CBIR organization [1] and [7].

In the CBIR, the visual substances of every image are retrieved from the image database. It is explained by multi-dimensional attribute vectors from which an image feature database is constructed [2] and [21]. Figure 1 displays the graphical presentation of the content-based image retrieval system. The consistency (similarity) between the query images and the images stored in the featured image files (database) is evaluated and the retrieval rate is executed by an image indexing process.

The problems in which the organization of data required, for such problems the unsupervised learning is used. Here, the structure detects the items based on a couple of distance measures (similarity) method [10]. The items that are related to one another are placed in one class (known as a
cluster) and the dissimilar items are put into different classes. The three visual traits that are color, shape, and textures are used in the combination of two and three image features for developing the following models: CS model (the combination of color and shape), ST model (the combination of shape and texture), CT model (the combination of color and texture), and CST model (the combination of all three traits values) [3].

Figure 1. The representation of Content-based image retrieval.

1.1. Some Existing CBIR Models:
Current methods of content-based image retrieval can be divided into two variations: the method of region-based image retrieval and the method of full-image retrieval. In full image retrieval methods, the features are taken out without segmenting into the regions of the entire image. So, the local features are used for the query image, and the global features are used for images in the image database [16]. In region-based methods, earlier to the extraction of the features the image is segmented into separate regions. Region-based methods can be further separated into three classes: In the first class, only the images in the image database are segmented and the query image is not segmented. In the second class, the images in the image database, as well as the query image, are segmented but only the single portion of the query image is used for the indexing/searching. The third class is the same as the second class only difference is that all the regions of the query image are to be used for the comparison [13]. Few of the existing CBIR methods: (i) Blobworld, developed by Carson et.al with the Computer Vision Group at UC Berkeley, is based on a region-based image retrieval technique. It segments the image into blobs using an EM-algorithm based on the color and texture traits of the pixels [4]. (ii) NaTra was developed at the University of California, Santa Barbara [14]. In this method, Images are usually segmented into six to twelve non-overlapping equal regions. (iii) PicSOM is a region-based image retrieval system developed at the Helsinki University of Technology's Computer and Information Science Laboratory. The separation of images in this system occurs in five areas. Color and texture characteristics are used for every area [11]. (iv) SIMPLicity (Semantics-sensitive Integrated Matching for Picture Libraries) was develop by Wang J Z et.al at Stanford University. It uses the k-mean clustering algorithm to part the image into regions [18]. (v) UFM (Unified Feature Matching) is a region-based image retrieval method developed by Chen and Wang in 2002 [5]. (vi) Chen et al. proposed in 2005, an unsupervised cluster-based image retrieval technique known as CLUE [6].

1.2. Organization of the paper
As follows, the remaining part of the paper is systemized. In Section2, the specifics of the research methodologies are elaborated. Section 3 addresses the findings and the interpretation of the results. In section4, the work is concluded.

2. Research methodologies
A few CBIR systems are developed by combining the color, shape, and texture of the three visual features. Individual features like color, shape, and texture of each image extracted by the well-defined method. Such as Color features are evaluated by color histogram and the color moment. The color
moments are calculated similarly to calculating moments of a probability distribution.
The principle moment of the color can be taken as the standard color in the picture and the following
mean equation can be used to evaluate it very well [8]:

\[ \text{Mean}_i = \frac{\sum_{j=1}^{N} P_{ij}}{N} \]  

(1)

Here, \( N \) determines the total number of possible pixels in the image and \( P_{ij} \) indicates the value \( i \)th
color channel and \( j \)th pixel of the image.
The second moment in colour is the Standard deviation (\( \sigma \)). It is achieved by evaluating the square
root of the variance of the colour distribution.

\[ \sigma = \sqrt{\left( \frac{1}{N} \sum_{j=1}^{N} 1 \ast |E_i - P_{ij}|^2 \right)} \]  

(2)

Here, \( E_i \) is the mean of the image \( i \)th color channel.
The skewness (\( S_i \)) is the third color moment. By the following equation, it can be calculated:

\[ S_i = \frac{1}{\sqrt{N}} \left( \frac{1}{N} \sum_{j=1}^{N} (P_{ij} - E_i)^3 \right) \]  

(3)

In terms of edge images computed using Gradient Vector Flow (GVF) fields, shape features are
considered and captured. It is mostly used in image processing for analyzing the number of block in
the image (retrieve shape feature of an image); it is introduced by Xu and Prince [19].
Let \( f(x, y) \) is an edge map classified on the image range. It is necessary for the regularity in results to
limit the edge map values to bounds between 0 and 1, and by virtue \( f(x, y) \) obtain on greater intensities
(near to 0) on the edges of the object. The GVF filed is known as the vector field and it is given by that
reduces the function of energy:

\[ \mathcal{H}(\nabla \mathcal{V}) = \iint_{\mathbb{R}^2} \left| \nabla \mathcal{V} \right|^2 + \mu (u x^2 + u y^2 + v x^2 + v y^2) dx \, dy \]  

(4)

Texture features are used to segment images and categorise those blobs into regions of interest. It is
characterized by the spatial sharing of intensity values in a region (blob) of the neighborhood. The
texture measure is defined by the following equation [9]:

\[ \text{Texture}(i, j) = \frac{1}{w^2} \sum_{m=-w}^{m+w} \sum_{n=-w}^{n+w} Edg(e(i + m, j + n)) \]  

(5)

Here, \( W = 2w + 1 \), it shows observation window size.
The auto-correlation method of an image can be utilized to distinguish redundant blueprints of
textures. The equation of the auto-correlation is given as [9]:

\[ \text{Autocorrelation}(p, q) = \sum_{i} \sum_{j} F(i, j) F(i - p, j - q) \]  

(6)

2.1. Proposed Model

By using the above equations, the CS, ST, CT, and CST CBIR models fuse the values of texture,
shape, and color features. These function values are then stored in the feature database. A limit of 0.7
(says a threshold value of 70%) is allotted for the texture, color and shape features values. The
mathematical model of the CS CBIR model is given by the combination of equations 1 and 4.
The mathematical model of the ST CBIR model is given by the combination of equations 4 and equation 5.

\[
ST = Shape(GVF) + Texture(i, j)
\]  

The mathematical model of the CT CBIR model is given by the combination of equations 1 and equation 6.

\[
CT = Si + Autocorrelation(p, q)
\]  

The mathematical model of the CST CBIR model is given by the combination of equations 1, 4, and 5.

\[
CST = Si + Shape(GVF) + Texture(i, j)
\]

The same limit is also set for the query image as 70 percent of the query image's color, shape, and texture function values. Compare the values of the target image's color, shape, and texture characteristics with the assigned threshold values. If the value of the color, shape, and texture visual characteristics of a target image exceeds the threshold value for the color, shape, and texture visual characteristics, the values of the color, shape, and texture visual characteristics of the target images are combined and stored in the database of stored features. If not, discard the target image as a suitable image [20].

2.2. Graph partitioning method

The visual feature values are usually calculated in advance and accumulated as visual feature value files for accumulated images in the database. Along with the image's similarity proportion, these function values are the similarity between the query and target images that are computed and processed afterwards. Next, as the neighbour of the query image, the quantities of target images that are extremely close to the query image are chosen as indicated by similarity measurements [15]. Figure 2 shows the neighbouring target image determination and groups of images that are the main part of the unsupervised CBIR process.

The initial step of execution of the CBIR method is to break the image into segments. For getting the image segment edges of the query (input) image should be detected. For achieving these following two steps must follow the first step, by applying the edge detection algorithm convert into the grey-scale. The second step is the extraction of features. By these two steps, an image is split into many clusters [22].

A set of X images is characterized by an undirected weighted graph \( G = (V, E) \), where V is the number of nodes signify images, i.e. \( V = \{1, 2, \ldots, n\} \), and the edges framed between each pair of nodes are E, i.e. \( E = \{(i, j): i, j \in V\} \).

Currently, the problem of graph partitioning can be formulated for clustering, using the CLUE spectral graph partitioning algorithm as the standardised partitioning technique for image clustering. A graph-partitioning technique is used to compose the similarity of the images (nodes) within the similarity of the group, as well as low comparability between the group clusters. It is likewise examined in [6].
Shi and Malik developed a model for solving the normalized cut of graph and image segmentation problems by using a generalized eigenvalue problem [16].

$$\begin{align*}
(D & \text{Diag} - w)y = \lambda(D \text{Diag})y \\
\text{Where } w & \text{ is a square affinity matrix of } n \times n, \text{ Diag= [a}_1, a_2, \ldots, a_n\text{] is a diagonal matrix with}
\end{align*}$$

$$a_i = \sum_{j=1}^{n} w_{ij}$$

In a certain graph illustration of images \(G = (V, E)\) with affinity matrix \(w\), suppose the number of image clusters be \(\{K_1, K_2, \ldots, K_m\}\), which is also the partition of \(V\), i.e., \(K_i \cap K_j = \emptyset\) for \(i \neq j\) and \(\bigcup_{i=1}^{m} K_i = V\).

Then the representative node (image) of \(K_i\) is

$$\text{max} \{\sum_{j \in K_i, k \notin K_i} w_{jk}\}$$

Here, \(w_{jk}\) is an affinity matrix value of \(i\)th and \(k\)th position, and \(K\) is the number of clusters, which can likewise be seen as a measure of the between clusters similarity.

2.3. Step by step description of the implementation

Stopping criteria: if a cluster has less than 100 images, then that cluster can’t be further clustered.

Total of 1000 images in the image database

STEP1: \(N = 1000\) (number of images in the database)

STEP2: Select one image at random (query image): group the neighbouring target images using the Nearest Neighbor Approach as per query image.

STEP3: First, construct an undirected weighted graph that uses equation 12 to include the query image as well as its neighbouring target images.

STEP4: Now the selection of a target matched node, the node has a maximum sum value of their feature select as a target matched in each cluster using equation 13.

STEP5: Apply N-partition to the graph for partitioning the graph into two subgraphs.

STEP6: Selection of Clusters: count the maximum number of images cluster and repeat
STEP4.

STEP7: If $N < 100$: further partition is not allowed

STEP8: Retrieval of Relevant Images: retrieve leaf clusters from left to right in the inorder traversal.

STEP9: Store each collection of images in a file for all queries (for all iterations) and count the precision at different levels of $k$, chosen $k = 10, 20, 30, \text{up to } 100$.

In the CBIR CS, ST, CT, and CST models, image feature values are stored in the stored feature files after adding visual feature colour, form, and texture values to an image with a minimum value of 70% of each feature in each image. The related images are extracted from the image database based on these weights of importance. Finally, the union of all four models is taken by normalizing the value between 0 and 1 shown in Figure 3.

![Figure 3. Union of four unsupervised content-based image retrieval models: CS, ST, CT, and CST.](image)

3. Results and Discussions

The experimental results have been performed with a universally useful COREL image database, which contained 10 distinct categories of images, each category contains 100 images of resolution 256 X 384, and a total approximately contained 1,000 images shown in Table 1 [23]. At present just the main 21 outcomes are introduced because of space restriction from all the CBIR models for one arbitrarily selected query image.

3.1. Results for bus class by all four models

Only for demonstrating purpose one query image from bus class (Image Id number 301) is chosen. Each image is treated in this work as a query image from each of the 10 groups of images and then an
aggregate of 1000 query images. To measure accuracy, the top k results were selected from the CBIR methods, i.e. known as precision at k. Results of each of the four models (CS, ST, CT, and CST) are shown in Figures 4 (a – d).

Table 1. COREL Image Database Depiction with Index Values [23]

| Class No. | Class Name       | Class No. | Class Name       |
|-----------|------------------|-----------|------------------|
| 1 (0-99)  | Village & People | 6 (500-599)| Elephants        |
| 2 (100-199)| Beach            | 7 (600-699)| Flowers          |
| 3 (200-299)| Buildings        | 8 (700-799)| Horses           |
| 4 (300-399)| Buses            | 9 (800-899)| Mountains and glaciers |
| 5 (400-499)| Dinosaurs        | 10 (900-999)| Food            |

CBIR system Results for bus class with the same query image.

Figure 4 (a). CS CBIR model result (the combination of shape and color features), the first image is an image of the query.

Figure 4 (b). ST CBIR model result (the combination of shape and texture features), the first image is an image of the query.

Figure 4 (c). CT CBIR model result (the combination of color and texture features), the first image is an image of the query.
There are 18 similar images are retrieved out of the top 21 results using the CS model from the bus class as shown in Figure 4(a), the first image is the query image. There are 19 similar images are retrieved using the ST model as shown in Figure 4(b). There are 20 similar images are retrieved using the CT model as shown in Figure 4(c). There are 21 out of 21 similar images are retrieved using the CST model as shown in Figure 4(d).

3.2. Performance evaluation
The average precision values of color-shape, color-texture, and color-shape-texture systems for every class at different precision levels (k = 10, 20, ….., 100) are computed. Here are taken 100 arbitrary queries of images. Precision considers all extracted images into account.

\[
\text{Precision} = \frac{|\text{Relevant Images upto } k|}{|\text{Total Images Retrieved upto } k|} 
\]

Let P is the total precision, P1, P2, P3, ………………… up to P100 are the precisions for image queries 1, 2, 3, ………………………..up to 100, of one particular class, because each class has a total of 100 images only. After calculating these precisions at level k, the average is taken and reported in the corresponding tables, where k is the different precision levels (k=10, 20, 30, ……, 100).

\[
P (a t \ k = 100) = \frac{|P1 + P2 + \ldots + P100|}{|k=100|} 
\]

In these experiments, the same feature extraction technique is used as given in [3]. The Euclidean distance is used as the measure of similarity to determine the similarity between the query and target images in the database [3].

\[
\text{Distance}(x, y) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} 
\]

3.3. Analyses of results
The performance at varying precision levels of k has been computed. The precision values of all four CBIR models (CS, ST, CT, and CST) for each class at the varying precision levels are given in Table 2, Table 3, Table 4, and Table 5 respectively, and the union values of all four models are given in Table 6. The results of the evaluation of the four CBIR systems at precision level 100 are given in Table 6.

The union of all four approaches has been performed by normalizing the incentive somewhere in the range of 0 and 1. The precision values at an average precision level 100 of all the CBIR approach with the union for each class of an image database are reported in Table 6. This is graphically represented in Figure 5. The average values of all categories at precision 100 of all four CBIR models with union graphically shown in Figure 6.
Table 2. Performance at different precision (k) of CS CBIR model (color and shape combination) with threshold of 0.7 for each class of image database.

| ID | Name  | 10  | 20  | 30  | 40  | 50  | 60  | 70  | 80  | 90  | 100 |
|----|-------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 1  | People| 0.70| 0.685| 0.667 | 0.636 | 0.615 | 0.595 | 0.59 | 0.585 | 0.58 | 0.57 |
| 2  | Beach | 0.68 | 0.645 | 0.617 | 0.586 | 0.554 | 0.535 | 0.505 | 0.475 | 0.446 | 0.42 |
| 3  | Buildings | 0.60 | 0.565 | 0.537 | 0.506 | 0.485 | 0.454 | 0.433 | 0.415 | 0.395 | 0.39 |
| 4  | Buses | 0.84 | 0.775 | 0.767 | 0.746 | 0.725 | 0.71  | 0.705 | 0.697 | 0.688 | 0.68 |
| 5  | Dinosaurs | 1.00 | 0.995 | 0.987 | 0.980 | 0.978 | 0.975 | 0.973 | 0.972 | 0.971 | 0.97 |
| 6  | Elephants | 0.58 | 0.535 | 0.487 | 0.436 | 0.395 | 0.386 | 0.375 | 0.375 | 0.369 | 0.36 |
| 7  | Flowers | 0.86 | 0.855 | 0.847 | 0.838 | 0.829 | 0.825 | 0.82  | 0.818 | 0.815 | 0.81 |
| 8  | Horses | 0.86 | 0.855 | 0.845 | 0.842 | 0.841 | 0.84  | 0.839 | 0.838 | 0.836 | 0.83 |
| 9  | Mountains | 0.54 | 0.525 | 0.517 | 0.492 | 0.475 | 0.442 | 0.415 | 0.398 | 0.387 | 0.37 |
| 10 | Food  | 0.78 | 0.765 | 0.742 | 0.727 | 0.713 | 0.702 | 0.7  | 0.697 | 0.695 | 0.69 |
| Avg All Categories | 0.744 | 0.72 | 0.701 | 0.678 | 0.661 | 0.646 | 0.636 | 0.627 | 0.618 | 0.618 | 0.609 |

Table 3. Performance of ST CBIR model in the combination of shape & texture features at varying precision levels of k with a threshold of 0.7 for each class of image database.

| ID | Name  | 10  | 20  | 30  | 40  | 50  | 60  | 70  | 80  | 90  | 100 |
|----|-------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 1  | People | 0.73 | 0.693 | 0.671 | 0.653 | 0.642 | 0.632 | 0.615 | 0.593 | 0.586 | 0.581 |
| 2  | Beach | 0.70 | 0.649 | 0.621 | 0.616 | 0.601 | 0.595 | 0.577 | 0.563 | 0.496 | 0.452 |
| 3  | Buildings | 0.64 | 0.572 | 0.559 | 0.536 | 0.523 | 0.511 | 0.486 | 0.471 | 0.423 | 0.413 |
| 4  | Buses | 0.83 | 0.779 | 0.771 | 0.763 | 0.753 | 0.742 | 0.737 | 0.727 | 0.713 | 0.712 |
| 5  | Dinosaurs | 1.00 | 0.996 | 0.992 | 0.987 | 0.981 | 0.976 | 0.975 | 0.974 | 0.973 | 0.971 |
| 6  | Elephants | 0.59 | 0.546 | 0.491 | 0.474 | 0.442 | 0.428 | 0.412 | 0.408 | 0.396 | 0.372 |
| 7  | Flowers | 0.88 | 0.858 | 0.853 | 0.842 | 0.831 | 0.828 | 0.823 | 0.820 | 0.817 | 0.813 |
| 8  | Horses | 0.87 | 0.858 | 0.851 | 0.853 | 0.848 | 0.844 | 0.841 | 0.840 | 0.838 | 0.832 |
| 9  | Mountains | 0.57 | 0.531 | 0.527 | 0.517 | 0.502 | 0.476 | 0.463 | 0.427 | 0.415 | 0.39 |
| 10 | Food  | 0.80 | 0.769 | 0.761 | 0.732 | 0.723 | 0.712 | 0.709 | 0.701 | 0.699 | 0.69 |
| Avg All Categories | 0.761 | 0.725 | 0.709 | 0.697 | 0.684 | 0.674 | 0.664 | 0.652 | 0.635 | 0.623 |

Table 4. Performance at different precision (k) of CT CBIR model (color and texture combination) with threshold of 0.7 for each class of image database.

| ID | Name  | 10  | 20  | 30  | 40  | 50  | 60  | 70  | 80  | 90  | 100 |
|----|-------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 1  | People| 0.70 | 0.685 | 0.676 | 0.646 | 0.625 | 0.615 | 0.598 | 0.587 | 0.582 | 0.57 |
| 2  | Beach | 0.68 | 0.645 | 0.628 | 0.605 | 0.575 | 0.555 | 0.525 | 0.485 | 0.476 | 0.46 |
| 3  | Buildings | 0.64 | 0.595 | 0.567 | 0.546 | 0.525 | 0.494 | 0.483 | 0.475 | 0.462 | 0.45 |
| 4  | Buses | 0.88 | 0.845 | 0.817 | 0.786 | 0.757 | 0.768 | 0.755 | 0.734 | 0.721 | 0.71 |
| 5  | Dinosaurs | 1.00 | 0.995 | 0.979 | 0.977 | 0.974 | 0.973 | 0.972 | 0.971 | 0.971 | 0.97 |
| 6  | Elephants | 0.58 | 0.565 | 0.556 | 0.526 | 0.495 | 0.487 | 0.477 | 0.454 | 0.444 | 0.43 |
| 7  | Flowers | 0.86 | 0.86 | 0.858 | 0.854 | 0.85 | 0.849 | 0.847 | 0.844 | 0.843 | 0.84 |
| 8  | Horses | 0.90 | 0.894 | 0.89 | 0.888 | 0.88 | 0.879 | 0.874 | 0.873 | 0.872 | 0.87 |
| 9  | Mountains | 0.55 | 0.53 | 0.513 | 0.505 | 0.492 | 0.471 | 0.465 | 0.454 | 0.442 | 0.43 |
| 10 | Food  | 0.78 | 0.77 | 0.769 | 0.767 | 0.761 | 0.754 | 0.75 | 0.749 | 0.742 | 0.74 |
| Avg All Categories | 0.757 | 0.738 | 0.725 | 0.72 | 0.695 | 0.684 | 0.674 | 0.662 | 0.655 | 0.647 |
Table 5. Performance at different precision (k) of CST CBIR model (color, shape & texture combination) with threshold of 0.7 for each class of image database.

| ID | Name     | 10  | 20  | 30  | 40  | 50  | 60  | 70  | 80  | 90  | 100 |
|----|----------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 1  | People   | 0.79| 0.785| 0.776| 0.746| 0.725| 0.715| 0.698| 0.687| 0.682| 0.65 |
| 2  | Beach    | 0.76| 0.745| 0.728| 0.705| 0.685| 0.655| 0.625| 0.589| 0.553| 0.52 |
| 3  | Buildings| 0.74| 0.725| 0.713| 0.686| 0.645| 0.614| 0.583| 0.547| 0.512| 0.49 |
| 4  | Buses    | 0.91| 0.89 | 0.872| 0.863| 0.815| 0.801| 0.795| 0.789| 0.78 | 0.77 |
| 5  | Dinosaurs| 1.00| 1.00 | 1.00 | 0.999| 0.997| 0.997| 0.995| 0.994| 0.992| 0.987|
| 6  | Elephants| 0.65| 0.643| 0.623| 0.603| 0.594| 0.576| 0.545| 0.528| 0.491| 0.47 |
| 7  | Flowers  | 0.94| 0.931| 0.928| 0.922| 0.915| 0.909| 0.903| 0.899| 0.895| 0.89 |
| 8  | Horses   | 0.96| 0.954| 0.951| 0.943| 0.937| 0.932| 0.929| 0.924| 0.92 | 0.92 |
| 9  | Mountains| 0.59| 0.578| 0.573| 0.565| 0.552| 0.549| 0.535| 0.517| 0.495| 0.46 |
| 10 | Food     | 0.84| 0.832| 0.826| 0.821| 0.809| 0.804| 0.799| 0.797| 0.793| 0.79 |
| Avg| All Categories | 0.818| 0.808| 0.799| 0.785| 0.767| 0.755| 0.74 | 0.727| 0.711| 0.694|

Table 6. Performance at an average precision 100 of all the CBIR approaches with the union for each class of image database.

| ID | Category Name | CS [19] | ST | CT [19] | CST | Union |
|----|---------------|---------|----|---------|-----|-------|
| 1  | People        | 0.57    | 0.581| 0.57    | 0.65| 0.72  |
| 2  | Beach         | 0.42    | 0.452| 0.46    | 0.52| 0.60  |
| 3  | Buildings     | 0.39    | 0.413| 0.45    | 0.49| 0.57  |
| 4  | Buses         | 0.68    | 0.712| 0.71    | 0.77| 0.83  |
| 5  | Dinosaurs     | 0.97    | 0.971| 0.97    | 0.987| 0.99  |
| 6  | Elephants     | 0.36    | 0.372| 0.43    | 0.47| 0.54  |
| 7  | Flowers       | 0.81    | 0.813| 0.84    | 0.89| 0.91  |
| 8  | Horses        | 0.83    | 0.832| 0.87    | 0.92| 0.97  |
| 9  | Mountains     | 0.37    | 0.39 | 0.43    | 0.46| 0.57  |
| 10 | Food          | 0.69    | 0.69 | 0.74    | 0.79| 0.88  |
| Avg| All Categories| 0.609   | 0.623| 0.647   | 0.694| 0.758 |

Figure 6 shows that a system based on a combination of three characteristics is stronger than a system based on a combination of two characteristics. Also, shows that the union of all four outperformed.

Figure 5. Results of Comparison of CS, ST, CT, CST, and union for each class of image database.
Figure 6. Average of all classes of different CBIR models (CS, ST, CT, and CST) with union.

4. Conclusion and future direction

Several CBIR models are developed with multiple features in a combination of two and three. The models with two features are designated as CS, CT, and ST models based on the combination of color with shape, color with texture, and shape with texture, respectively. The three features based model is developed using color, shape, and texture labeled as CST model. The comparative performance of these models for image retrieval is assessed using the open-source image database COREL. With the data subset considered, it is found that among two features-based models, the CT model gives the best average performance of 64.7%. It is also observed that combining three features in the CST model results in further improvement in accuracy of image retrieval up to 69.4%.

In the future, other CBIR models may be proposed with additional spatial features as well as with more number of features in combination to check for possible improvement of retrieval accuracy.

References

[1] Ahmed, Khawaja et.al. “Content Based Image Retrieval using Image Features Information Fusion”, Journal of Information Fusion. Vol. 51, pp. 76 – 99, 2019.
[2] A. K. Alhassan and A. A. Alfaki, "Color and texture fusion-based method for content-based Image Retrieval", International Conference on Communication, Control, Computing and Electronics Engineering, pp. 1-6, 2017.
[3] M. K. Alsmadi, “Content-Based Image Retrieval Using Color, Shape and Texture Descriptors and Features”, Arabian Journal for Science and Engineering, Vol. 45, pp. 3317–3330, 2020.
[4] C. Carson, et.al., “Blobworld: Image Segmentation Using Expectation-Maximization and its Application to Image Querying”, IEEE Trans. Pattern Anal. Machine Intell., vol. 24, no. 8, pp. 1026–1038, 2002.
[5] Y. Chen and J. Z. Wang, “A Region-Based Fuzzy Feature Matching Approach to Content-Based Image Retrieval”, IEEE Transaction on Pattern Analysis and Machine Intelligence, vol. 24, no. 9, pp. 1252-1267, September 2002.
[6] Y. Chen, J. Z. Wang, and Robert krovetz, “CLUE: Cluster-Based Retrieval of Images by Unsupervised learning”, IEEE Transaction on Image Processing, Vol. 14, No. 8, pp. 1187-1201, 2005.
[7] R. Datta, et.al., “Image retrieval: Ideas, influences, and trends of the new age”, ACM Comput. Surv., Vol. 40, no. 2, Article 5, 60 pages, 2008.
[8] J. Han, and K. Kuang, “Fuzzy color histogram and its use in color image retrieval”, IEEE Trans. Image Pro cessing, Vol. 11, No. 8, pp. 44 – 952, 2000.
[9] R. M. Haralick, et.al., "Textural Features for Image Classification", IEEE Transactions on Systems, Man, and Cybernetics, SMC-3 (6): 610–621, 1973.
[10] S. Jabeen, et.al., “An effective content-based image retrieval technique for image visuals representation based on the bag-of-visual-words model”, PLoS ONE 13(4), 2018.
[11] J. Laaksonen, et.al., “Picsom - Content-based image retrieval with Self-Organizing Maps”, Pattern Recognition Letters, vol. 21, pp.1199–1207, 2000.
[12] J. Li and J. Z. Wang, “Real-time computerized annotation of pictures”, IEEE Trans. Pattern Analysis Machine
[13] Y. Liu, et.al., “A survey of content-based image retrieval with high-level semantics”, Pattern Recognition, Volume 40, Issue 1, pp. 262-282, 2007.
[14] W.Y. Ma and B. Manjunath, “NaTra: A Toolbox for Navigating Large Image Databases,” Proc. IEEE International Conference on Image Processing, pp. 568-571, vol. 8, no. 20, 1997.
[15] Z. Mehmood, et.al., "Content-Based Image Retrieval Based on Visual Words Fusion Versus Features Fusion of Local and Global Features", Arab J Sci Eng, Vol. 43, pp. 7265–7284, 2018.
[16] J. Shi and J. Malik, “Normalized Cuts and Image Segmentation”, IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 22, No. 8, pp. 888 – 905, 2000.
[17] A. Smeulders, et.al., “Content-based image retrieval at the end of the early years”, IEEE Transaction Pattern Analysis Machine Intelligence, vol. 22, No. 12, pp. 1349–1380, 2000.
[18] J. Z. Wang, J. Li, and G. Wiederhold, “SIMPLicity: Semantics-Sensitive Integrated Matching for Picture Libraries”, IEEE Trans. Pattern Anal. Mach. Intell., vol. 23, no. 9, pp. 947-963, 2001.
[19] C. Xu and J.L. Prince, "Snakes, Shapes, and Gradient Vector Flow", IEEE Transactions on Image Processing. 7 (3): 359–369, 1998.
[20] S. M. Zakariya, et.al., "Unsupervised Content Based Image Retrieval by Combining Visual Features of an Image With A Threshold," International Journal of Computer and Communication Technology, Vol. 3: Iss. 3, 2012.
[21] S. M. Zakariya, et.al., “Combining visual features of an image at different precision value of unsupervised content based image retrieval,” 2010 IEEE ICCICR, pp. 1 – 4, 2010.
[22] S. Unar, et.al., “A decisive content based image retrieval approach for feature fusion in visual and textual images”, Journal of Knowledge-Based Systems, Vol. 179, pp. 8 – 20, 2019.
[23] COREL Image Database is obtained from: http://wang.ist.psu.edu/docs/related/