Region-based image retrieval using region of interest (ROI) according to incremental frame and clustering color image

Abd Rasid Mamat *, Fatma Susilawati Mohamad, Mohamed Afendee Mohamed, Norkhairani Abd Rawi, Mohd Khalid Awang

Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin, Kuala Terengganu, Malaysia

ARTICLE INFO

ABSTRACT

Content-based image retrieval involves the extraction of global feature images for their retrieval performance in large image databases. Extraction of global features image cause problem of the semantic gap between the high-level meaning and low-level visual features images. In this study RBIR, Region of Interest Based (ROI) Image Retrieval Using Incremental Frame of Color Image was proposed. It combines several methods, including filtering process, image partitioning using clustering and incremental frame formation, complementation law of theory set to generate ROI, NROI, or ER of the region. The concept of weighting as well as a significant query is also incorporated as a query strategy. Extensive experiments were also conducted on the Wang database and the color model selected was the CIE lab. Experimental results show the proposed method is efficient in image retrieval. The performance of the proposed method shows a better average IPR value of 3.51% compared to RGB and 22.92% with the HSV color model. Meanwhile, it also performs better by 36%, 5%, and 24% compared to methods CH (8,2,2), CH (8,3,3), and CH (16,4,4).

1. Introduction

Today, technological developments have seen the rapid growth of the Internet, covering wider areas, increasing the usage of multimedia and electronic devices equipped with advanced technology at affordable prices, enabling millions of images to be uploaded and downloaded in the media social led to the growth of the digital image database exponentially. This triggered researcher to come out with efficient image retrieval techniques with the aim of finding the images of interest. This opportunity has been taken by CBIR as a new approach to efficiently retrieve relevant images. This technique tends to retrieve similar images using visual features to a given query.

Despite the many publications using visual features in CBIR to report the performance and significant advances in image retrieval, this technique still shows some limitations. Extraction of visual features from the whole image, causing a problem of the semantic gap between the high-level meaning and low-level visual features images. These visuals insufficiently describe some relevant objects or the particular region that the user is interested in (Bchir et al., 2018). Another limitation is their sensitivity to the type the visual features that are extracted from the images. Absolutely, the visual feature relevance depends on the image content. Therefore, the more accurate the retrieval result obtained, the more distinctive the visual features are. Actually, achieving efficient image retrieval can be done by allowing the user to select the important or desired specific region that will minimize the semantic gap.

In this paper, we propose a region of interest (ROI) based on image retrieval. To obtain the ROI of images, the clustering process, dividing non-overlapping images, the concept of complementary law of basic set theory is wrapped up as a part of the proposed method. The silhouette index is also consolidated to validate the resulting clusters. The rest of the paper is organized as follows: Section 2 represents a related work for this research. Section 3 describes the proposed technique that triggers the idea resulting from chapter 2. This section also covers the process of measuring the similarity and performance of the ROI-based image retrieval. Results and discussion is presented in section 4 and

* Corresponding Author.
Email Address: arm@unisza.edu.my (A. R. Mamat)
https://doi.org/10.21833/ijaas.2022.01.016
Corresponding author's ORCID profile:
https://orcid.org/0000-0003-0047-0945

finally section 5 elaborates the conclusion of this research.

2. Related works

Region of Interest (ROI) or Object-based image retrieval (RBIR/OBIR) emerged as an alternative approach to effectively retrieve relevant regions or objects in images that the user may be interested in. This approach attracted many researchers for expanding their research and building commercial products (Vu et al., 2003; Chan et al., 2008; Moghaddam et al., 2001; Tian et al., 2000; Zhou et al., 2005; Lee et al., 2012; Vimina and Jacob, 2013; Eze et al., 2019; Baji and Mocanu, 2017; Raja et al., 2020). This method offers a way to reduce the semantic gap between the visual properties of the query and the high-level understanding of the user. Kam et al. (2000) mentioned the performance of effective CBIR lies in the ability to access the image at the level of objects. This is because users typically want to search images that contain important or particular objects of interest. Therefore, for the proposed RBIR from CBIR where the representation, index, and query level using the object is critical (Carson et al., 1999). The authors proposed RBIR based on object extraction through image segmentation. A multiscale segmentation algorithm that automates the segmentation process was applied in this research, the output of which is assigned novel color and texture descriptors that are both efficient and effective. To improve the result, the authors perform strategies of the query.

Next, Tian et al. (2000) proposed a novel method of CBIR by combining user-defined Region-of-Interest (ROI) and spatial layout. The image is partitioned as n x n non-overlapping image blocks and features such as color and texture are extracted from each block. Selections of 2 x 2, 3 x 3, 4 x 4 and 5 x 5 layouts are available depending on the complexity of the internal structures of the query image. Then the spatial layout is compared to the conventional approach using global features in the relevance feedback process. Next, user-defined ROI is applied as a refined spatial layout approach. Since the automatic segmentation to perform ROI is not always reliable, the authors suggested to capture of image object (ROI) is done by the user. Therefore, more accurate relevance feedback is achieved and thus leads to a better result.

The combination of image features based on color, shape, and location to perform region matching for image retrieval was proposed by Prasad et al. (2004). Dominant regions within each image are indexed using integrated color, shape, and location features. Various combinations of regions are also indexed. To obtain individual regions, images are divided into 3×3 blocks, and the number of blocks with the largest area of the region is designated. The hash structure is used to store the resulting index and related metadata. The retrieval process is non-cascading and images can be retrieved based on color, shape, or location and also based on a combined color-shape-location index. The results portray the effectiveness of retrieval increases in non-cascaded region-based querying by combined index.

Kim et al. (2004) proposed a new automatic method to classify either object or non-object of image. A set of regions located near the center of the image is defined as an object. It is because the color distribution is significant or more noticeable near the center of the image (object) compared to the surrounding area or background of the objects. Three measures for the classification are based on the characteristics of an object. Firstly, calculate the center significance based on the different color distribution between the areas of the center and the surrounding region. Secondly, compute the variance of significantly correlated colors in the image plane. Significantly correlated colors are first defined as the colors of two adjacent pixels that appear more frequently around the center of an image rather than at the background of the image. Finally, to estimate the object based on the edge strength at the boundary of the region. A combination of these measures is used to classify the object of the image by training the neural network.

Meanwhile, Zhou et al. (2005) presented a new method for CBIR that combines the detection of the ROI with relevant feedback. To describe the image contains the ROI approach more accurately compare using global features, meanwhile the relevance feedback makes the system adaptable to the subjective human perception. Color, texture, and shape features are extracted from the ROI. For an illustration of the overall approach, the authors apply color saliency and wavelet feature saliency to determine the ROI. The result portrays that the proposed method performs better performance compared to the global feature-based approaches and region-based techniques without feedback. Next, Chan et al. (2008) presented a novel image feature called color variances among adjacent objects (CVAAO). The CVAAO is not sensitive to distortion and scale variations of images. For contiguous objects in an image, it can effectively describe the principal colors and texture distribution of the image. According to CVAAO, the authors proposed two methods, using full image and clip or region of the image. CVAAO-based image retrieval method for full image meanwhile CVAAO-based on ROI for clip or region image. The results show that the CVAAO-based ROI image retrieval method better performance in finding out the database images that meet user requirements.

Wang et al. (2013) proposed an ROI-based image retrieval. This research considers user-defined the ROI query more effectively compare to ROI queries automatically in terms of the user's intention. Two characteristics appear when used to define the ROI, firstly the target region is located at the center of the ROI query and the ROI query contains hardly any noisy descriptors which do not belong to the target region. The proposed system integrated the two characteristics above, general bag-of-words image

retrieval method and an auxiliary Gaussian weighting (AGW) scheme. A 2-D Gaussian window function is used to weigh each descriptor according to its distance between the center of the ROI query. The score of each image database is computed using the AGW scheme and an efficient re-ranking algorithm is proposed based on the distribution consistency of the Gaussian weight between the matched descriptors of the ROI query and the candidate image. The results demonstrate the proposed system obtain satisfactory retrieval results.

Yang et al. (2011) proposed object retrieval to improve performance. The system exploits the information about the visual context of the query object and employs it to compensate for possible uncertainty in feature-based query object representation. Visual elements surrounding the query object in the query image are known as Contextual information. The authors also consider the region of interest (ROI) based on, firstly an uncertain observation of the latent search intent and second the saliency map detected for the query image. Then the search intent scores are integrated into the contextual object retrieval COR model to more effectively meet users' true information needs. Several experiments are conducted on several datasets and the results demonstrate some improvements in the object retrieval performance.

The discussion of some state of art techniques for ROI image retrieval was presented by Shrivastava and Tyagi (2015). In this research, the authors mentioned the problems faced and summarized solutions provided by various researchers for ROI image retrieval and then performed the generalized overall framework for ROI image retrieval. Based on the study it was found that there are many problems related to ROI image retrieval which have not been answered satisfactorily till now. Among them are, first, the accurately reflecting the user intent in query formulation, second, an effective technique for selecting ROI overlapping blocks, third, the technique for considering relative locations of multiple ROIs and finally, the reducing of overall computation time for region matching without affecting the accuracy of the system.

The RBIR system is based on color, texture, and shape features proposed by Vimina and Jacob (2013). Local color, texture, and shape features are extracted from selected sub-blocks of images, and the global color is extracted from the whole mages. All images are divided into different sized blocks ((3x3 grid/blocks, horizontal and vertical grids/block, central block, and the entire image) for extracted features. Histograms of the quantized HSV color space represent color features, texture features namely contrast, energy, correlation, and homogeneity from Gray Level Co-Occurrence Matrix (GLCM) represent the texture and Edge Histogram Descriptor (EHD) features represent shape features. For finding the minimum distance between the sub-blocks of the query and target image, the authors also used modified Integrated Region Matching

(IRM) algorithm, and the proposed method showed better performance compare to some previous methods.

A combination of the mean shift tracking (MST) and an improved expectation-maximization (EM)-like (IEML) methods have been used by Chen et al. (2015) to produce a novel ROI image retrieval. MST functions to seek the initial location of the target candidate model meanwhile IEML is used to adaptively change the location and scale of the target candidate model. This function aims to include the relevant region and exclude the irrelevant region as far as possible. As part of this method, images are divided into 4 blocks/regions and used color histogram and spatial features. Compared with the previous, the proposed method can directly find the target candidate model in the candidate image without pre-segmentation in advance and thus causes some performance improvement.

Eze et al. (2019) emphasized that content authentication and verification in medical images are required to obtain the region of interest (ROI) in RBIR. This negligence can lead to adversarial modification of the stored image which could have a lethal effect on research, diagnostic outcome, and the outcome of some forensic investigations. The proposed method combined both robust watermarking and fragile steganography with image search features to design a medical RBIR system that incorporates ROI integrity verification. Original ROI features were pre-computed and embedded into archival images and utilized during retrieval for image integrity checks.

Next, Baji and Mocanu (2017) proposed a new method for RBIR using the region of interest (ROI) of images. The ROI technique is based on segmenting the image into fixed partitions. This method is also based on the connected components and interesting objects to generate the histogram and statistical texture feature vectors. Color and texture features of the connected components are computed from the histograms of the quantized HSV color space and Gray Level Co-occurrence Matrix (GLCM), respectively. Histogram intersection is used for the matching process. From the experiments, without any knowledge, the proposed method was able to extract interesting objects from uninteresting objects and a complex background image.

Raja et al. (2020) proposed a new method for RBIR using color features. In this method, the authors used Sobel and canny technique to find the region of interest (ROI) with HSV color space was used. For classification, a neural network is used having categorized the data with class labels. To compute the similarity difference, techniques such as like Manhattan distance, Euclidean distance, Chebyshev, Hamming distance, and Jaccard distance are used. The proposed method improves the accuracy and precision result.

From our study it was found that using ROI we can reduce the semantic gap between visual properties and high-level human understanding of the images. However, there is still space in ROI for

use in a more efficient image retrieval system. This is a challenges researcher need to face in improving the performance of image retrieval systems using ROI techniques.

3. Proposed method

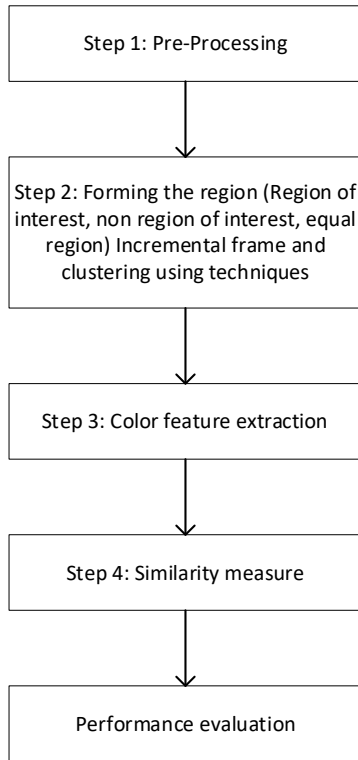Fig. 1 represents the block diagram for the proposed method.

Fig. 1: The block diagram for the proposed method

3.1. Step 1: Pre-processing

This phase involved two processes namely data acquisition and filtering the images. Data acquisition is obtained from an image from the Wang database (Yue et al., 2011; Mamat et al., 2016a). This database contains 1000 images and is divided into 10 categories and is known as African People, Beach, Building, Buses, Dinosaurs, Elephants, Flowers Horses, Mountains, and Food.

Meanwhile, the filtering process is performed to reduce the noise and to increase the quality of the image (Singh and Hemachandran, 2012; Hossain and Islam, 2017; Mamat et al., 2015). To attain this aim, we apply the median filter to an images Median filter is used because this filtering is performing better than the average filtering in the sense of removing impulse noise (Manoharan and Sathappan, 2013; Kannan et al., 2010; Szeliski, 2010; Malviya et al., 2017). Median filter filters that output at each voxel (i, j, k) the median of density values of an input image in the neighborhood of (i, j, k). A voxel is a unit of graphic information that defines a point in three-dimensional space ($x, y,$ and z coordinates). The definition of Median filter is as follows (Toriwaki and Yoshida, 2009):

$$Median\ Filter\ (F): F = \{f_{ijk}\} \to G = \{g_{ijk}\}$$

$$\{g_{ijk}\} = the\ median\ of\ \{f_{pqr;}(p,q,r) \in N((i,j,k) \qquad (1)$$

where $N((i,j,k))$ = the neighbourhood.

3.2. Step 2: Forming the region

Performing the region is an important part of this paper. The output of this step is identified the region is as the region interest region (ROI), non-region of interest (NROI), or equal region (ER). Several proses are integrated such as incremental frames, k-means clustering, cluster validation, basic theory set-complement law.

i) Incremental Frame: Portioning the images using an incremental frame. It's stated from frame 1 and increment to frame 5. Using this Frame, allows the entire image area to be used. This gives the advantage that an important object/area is concentrated somewhere in a smaller frame so it has a greater weight value. Similarly, if an important object/area bursts a lot of space in a larger frame then it will gain more weight. This situation gives the advantage of the database being used as it generally contains various positions of objects in the image, from simple images (a small number of objects) to more complex images.

An example of an incremental Frame is shown in Fig. 2. The steps to get this frame are as follows:

1. Get the midpoint pixel of the image
2. Move the pixels from this center point left, right, up, and down to form a rectangular frame (F).
3. If the frame size centered on this midpoint forms an area of as much as:

- 20% of the total, then the frame is marked as F_1,
- 40% of the total, then the frame is marked as F_2,
- 60% of the total, then the frame is marked as F_3,
- 80% of the total, then the frame is marked as F_4,
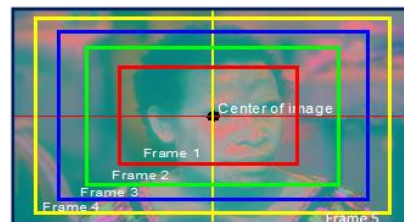- 100% of the total, then the frame is marked as F_5.

Fig. 2: An example of the incremental frame for image portioning

ii) K-Means Clustering: To obtain a cluster of regions, k-mean clustering was used. It is because is one of the most widely used clusterings and is the ease of implementation, simplicity, efficiency, and empirical success (Mamat et al., 2015; Zambre and Patil, 2013; Rejito et al., 2017). Fig. 3 shows an example of clusters resulting in k-means. In this example k=3.

Fig. 3: An example of k-means clustering

iii) Cluster Validation-Silhouette Index: The aim of cluster validation is to evaluate the goodness of clusters that produce (Maulik and Bandyopadhyay, 2002) and it is one of the major concerns and essentially important to the success of clustering applications (Jain and Dubes, 1988). Silhouette Index (SI) is a well-known technique for cluster validation (Chaimontree et al., 2010). Besides, according to Arbelaitz et al. (2013), in their research, this index portrays one of the best performing measurements. It is capable of pointing out which objects were placed well within their cluster and which ones are merely somewhere in between clusters. In this research, the authors use only *k*=2 because to take advantage of the goodness of the cluster (Mamat et al., 2018).

iv) Complementation Law-Set Theory: The definition of Complementation Law of set theory. If U is a universal set and A be any subset of U then the complement of A is the set of all members of the universal set U which are not the elements of A (Talib, 2006).

$$A' = \{x: x \in U \text{ and } x \notin A\} \qquad (2)$$

Adaptation of this theory is also used as a part to perform ROI or vice versa. In this case, U=Frame (F), caused by k=2 where only two clusters are in *F* and x are pixels that form either cluster 1 or 2*.*

If $A > A'$, A is ROI and A' is NROI in F.
If $A = A'$, A and A' is ER in F.

3.3. Step 3: Colour features extraction

Color is one of the most extensively used visual content for image retrieval. This is because it is easy to extract rather than shape and texture. In addition, the color feature is relatively robust to background complication and independent of image size and orientation and this has attracted many researchers to use it in their research (Mamat et al., 2016a; Singh and Hemachandran, 2012; Hossain and Islam, 2017; Mamat et al., 2015). Taking advantage of these, color features were used in this study. It includes color moment (CM) and manipulation of color histogram consisting of the color histogram (H), normalized color histogram (NH), and bin of color histogram (BH). The equation below is used to calculate the color features:

i) Color Moment: The first three moments, mean, variance and standard deviation have shown to be efficient and effective in representing color

distributions of images (Mamat et al., 2015; Chaudhuri, 2007). The equation below is used to compute the mean, variance, and standard deviation of an image of size N x M.

$$\text{Mean } (\bar{x}) = \frac{\sum_{i=1}^{N}\sum_{j=1}^{M} x_{ij}}{NM} \qquad (3)$$

$$\text{Variance } (\delta^2) = \frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}(x_{ij} - \bar{x})^2 \qquad (4)$$

$$\text{Standard deviation } (\delta) = \sqrt{\frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M}(x_{ij} - \bar{x})^2} \qquad (5)$$

where x_{ij} is the value of the pixel in row *i* and column *j.*

ii) Color Histogram (CH): The histogram *h* for a color image *I* (separated to individual color channel) with intensity value in the range *I (u, v)* ∈ [0, *K-1*], where *K*=256. Every single histogram is defined as in Burger and Burge (2016).

h(i)=the number of pixels in *I* (separated to individual color channel) with the intensity value *I*, for all 0≤*i*<*K* and more formally such as:

$$h(i) = \text{card } \{(u, v)| I (u,v)=i\} \qquad (6)$$

iii) Normalized Color Histogram (NCH): The normalized color histogram H' is given as (Burger and Burge, 2016):

$$H'=\{H' [0], H' [1], H' [2], H'...H' [i],, H'[n] \} \qquad (7)$$

where H'=(H [i])/p and p is the total number of pixels of an image.

iv) Bin Color Histogram (BCH): Color histogram can represent a range of intensity values and this technique is referred to as bin or bucket. This bin contains the collection of a range of pixel values. In a binned color histogram of size B, each bin h (*j*) contains the number of image elements/pixels having values within the interval $[a_j, a_{j+1}]$. Therefore, the definition color histogram bin such that of Szeliski (2010) is as follow:

$$h(j) = card \{ (u,v)|a_j \le I (u,v) < a_{j+1}\} \qquad (8)$$

for 0≤j<*B*. The range of possible values in *B* is divided into bins of equal size (bin width) k_B=K/B such that the starting value of *j* is:

$$a_j = j. \frac{K}{B} = j. k_B \qquad (9)$$

3.4. Step 4: Similarity measure

To choose the method for determining the similarity or distance of the images depends solely on the CBIR choice. The similarity tries to capture the strength of the relationship between features during comparisons of images in a database. Many types of techniques to compute the distance such as City Block, Manhattan, Euclidean distance (ED), and many others. Because of widely and ease to use, this

study applies ED for the RBIR system (Zhang and Lu, 2002).

$$ED=\sqrt{\alpha(ROI)^2 + \beta(NROI)^2}$$
$$=\sqrt{\alpha\{(CM_q - CM_i)^2 + (H_q - H_i)^2 + (NH_q - NH_i)^2 + ..}$$
$$=\sqrt{..\alpha\{(BH_q - BH_i)^2\} + \beta\{(CM_q - CM_i)^2 +}$$
$$=\sqrt{..+(H_q - H_i)^2 + (NH_q - NH_i)^2\}} \qquad (10)$$

where *q*=query, *i*=index (image in the database), $\alpha+\beta$=size of the frame. Next:

α=weight for ROI, which is $\alpha=\alpha/(\alpha+\beta)$ in condition $\alpha>\beta$ and
β=weight for NROI, which is $\beta=\beta/(\beta+\alpha)$ in condition $\beta<\alpha$ and,
if $\alpha=\beta$ weight for ER, which is $\alpha/(\alpha+\beta)=\beta/(\beta+\alpha)$

3.5. Step 5: Performance evaluation

ROI-based image retrieval aims for searching image databases for specific images that are similar to a given query image. There are many RBIR systems available but it is difficult to determine which is the best ones because it is impossible to compare RBIR systems quantitatively and objectively. One reason is the absence of a standard database to determine a set of quantitative performance measures. Two commonly used measurements of performance are precision and recall.

i) Precision and Recall: Performance of ROI system is evaluated and analyzed through Precision and Recall (Liu et al., 2007; Mamat et al., 2016b). Precision (*P*) is defined as the ratio of the number of relevant images retrieved (N_r) to the number of total of the images retrieved *K,* whilst Recall (*R*) is defined as the number of retrieved relevant images N_r, over the total number of relevant images available in the database N_t. Precision and Recall is calculated using Eqs. 11 and 12.

$$Recall = \frac{N_r}{N_t} \qquad (11)$$
$$Precision = \frac{N_r}{K} \qquad (12)$$

ii) Interpolated Precision and Recall (IPR): Interpolated Precision and Recall is related to precision and recall. One of the goals of IPR is to produce a smoother precision and recall graph. Precision and recall graph has a distinctive sawtooth shape, example if the $(K+1)^{th}$ image retrieved is a non-relevant image and the value of recall is the same as for the top *K* images. In this case, the value of precision has dropped.

To overcome this problem the interpolated precision is used (Manning et al., 2009; Keilwagen et al., 2014) The IPR can be expressed mathematically as Eq. 13 below, and calculating the value of R, will give the same value as IR. Fig. 4 represent precision and recall graph before and after IPR.
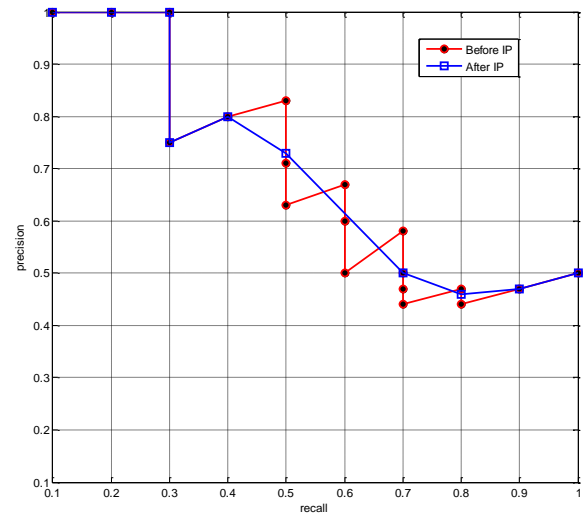
$$IPR = PR_1 + PR_2 + \cdots + PR_n/n \qquad (13)$$

Fig. 4: Example of precision and recall graph before and after IP (IPR)

iii) Significant Query (SQ) using Average precision: One of the strategies to improve the performance of ROI-based systems is suggested by the use of significant queries (SQ). 10% of images (10 queries) per category are used as queries. The steps to get an SQ are as follows:

a) Obtain the threshold, $\vartheta_{t,c}$ for interpolated precision (IP) for a particular image query (eg 10 return images for the first 10 queries) for each category. In this study $\vartheta_{t,c}$ is the average IP and *t* is a variable depending on the IP number of the query. $\vartheta_{t,c}$ is calculated using Eq. 14.

$$\vartheta_{t,c} = \sum_{i=1,2,..t}^{t}\left(\frac{IP}{t}\right)_c \qquad (14)$$

where *c*=category of image.

b) Compare IP query (Q) with values $\vartheta_{t,c}$ and if IP query is more than and an equal=$\vartheta_{t,c}$ then Q=SQ and vice versa.

iv) Experimental setup: Database was developed by Wang and his colleagues from the Pennsylvania State University (Li and Wang, 2003) and is available at http://wang.ist.psu/edu is used for this research.

- Experiment 1: Aim of this experiment is to compute Precision and Recall (PR) and follow by IPR. 10% of the images are used as query and each query consist of 5 frames (*F*) namely*F1*(F_{20}) *F2* (F_{40}), *F3*(F_{60}),*F4* (F_{80}) and*F5* (F_{100}). Therefore, each image will produce 5 queries (1 image x 5 frames), and each category that consists of 10 query image produce 50 queries (10 image x 5 frames). Next, get Average Interpolated Precision values for all queries per frame for each category.

Use the SQ method in the previous subsection to obtain the SQ.

- Experiment 2: Repeat experiment 1 using SQ and then calculate Interpolated Precisian Recall (IPR) SQ image of each frame for each category. Next, get an average IPR of SQ queries for all frames of each category. Figs. 5 and 6 show the average IPR for Frame 1 and Frame 2 for category 10 for SQ. Axis x=Recall=Average IR and y axis=Accuracy=Average IP.
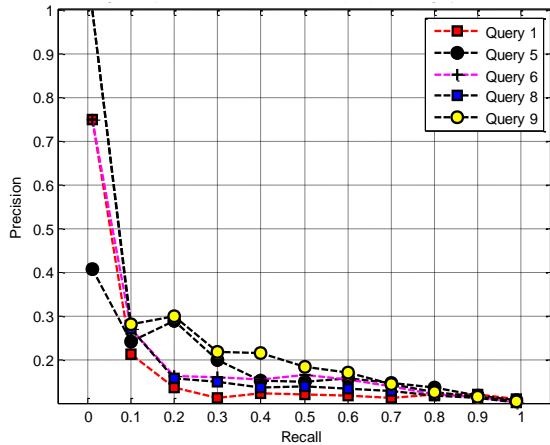

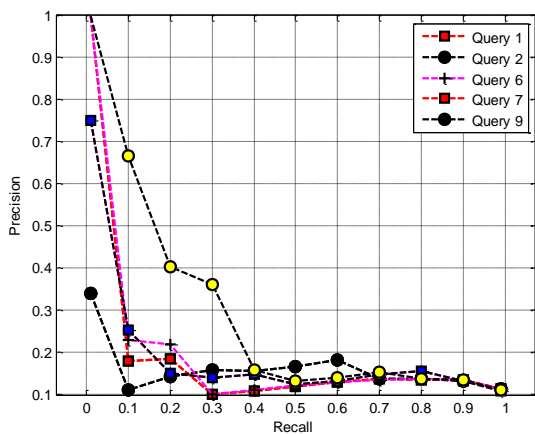**Fig. 5:** Average IPR for Frame 20 of SQ (q1, q5, q6 q8, and q9)

Fig. 6: Average IPR for Frame 40 of SQ (q1, q2, q6 q7, and q9)

4. Results and discussion

This section discusses the experimental results that have been performed and are divided into two sections as in the paragraph below.

4.1. Comparison between frames for each category

This comparison aims to find out which frame performs best for each category and the result is shown in Table 1. The result is based on the Average IPR of 20 return images (Can use any number of return image variables. This data is selected as it will be used for comparison with other researchers). In general, one of the reasons for this result is that the position of the homogenous object/region is scattered unevenly within the frame. As an example

for the category 5-Dinosaur, the best performing frame is F2, meaning that the region of interest lies in the middle and the surrounding area and the opposite results apply to categories 9 and 10.

Table 1: The best frame of each category for 20 return images

Category	Frame
1-Africa	$F3(F_{60})$
2-Beach	$F5(F_{100})$
3-Building	$F4(F_{80})$
4-Bus	$F5(F_{100})$
5-Dinosaur	$F2(F_{40})$
6-Elephant	$F5(F_{100})$
7-Flower	$F4(F_{80})$
8-Horse	$F3(F_{60})$
9-Mountain	$F5(F_{100})$
10-Food	$F5(F_{100})$

4.2. Comparison with other research

This section compares the findings with other researchers (Ayan et al., 2016). The comparisons have been shown in Table 2 and Table 3 below. In Table 2, comparisons are made with the best performance in the RGB color model and HSV color model. For RGB color models, the performance of the proposed method is better for Category 1, 3, 4, 7 9, and 10, while in the HSV color model, the category that outperforms is Category 1 2, 3, 4 7, 9, and 10. The proposed method shows a better average IPR value of 3.51% compared to RGB and 22.92% with the HSV color model.

Next compare the performance of the proposed method with the performance of CH (8,2,2), CH (8,3,3) CH (16,4,4), and the best of CH combinations as shown in Table 3. It is found that the proposed method is better than CH (8,2,2) for the 1,2, 3, 6,7, 9, and 10 categories, and for CH (8,3,3) the best performance of the proposed method is the categories 2,3,7,9 10 and the same for category 6. Meanwhile, for CH (16,4,4), the proposed method is shown better in categories 2,6,7,9,10 and the same value for categories 3 and 6.

5. Conclusion

In this paper, ROI based incremental frame for image retrieval is presented to enhance the retrieval performance. The various techniques used to include the filtering process, image partitioning using clustering and incremental frame formation, complementation law (theory set) have been blended to construct the proposed method. The results showing the best performance of frames are not static for this database. In general, the ROI position in the image is scattered. Some are located in the center of the image (e.g., Dinosaur-F5) and some are dispersed throughout the image (e.g., Mountain-F9).

The use of ROI in image retrieval is effective for ROI dispersed in whole images (examples: F2-Beach, F7-Flower, F9-Mountain, F10-Food) and vice versa (example F5-Dinosaurs). However, there are other factors that influence the result, for example, the

nature of objects in images, complexity (such as single or multiples objects and variety of color) images, limitation of the proposed method, and others. Compared to other researchers, the proposed method is better performance is 36%, 5%, and 24% compared to methods CH (8, 2, 2), CH (8, 3, 3), and CH (16, 4, 4)). In addition, we plan to use other databases that contain fewer object and shape features to evaluate the proposed method.

Table 2: Comparison of performance of precision for results of 20 image retrievals for all categories in RGB and HSV color model with the proposed method

Category	Ayan et al. (2016)		Proposed Method	Performance Differences	
	Best performance of precision (RGB Color Model)	Best performance of precision (HSV Color Model)	Best performance of Precision (CIE Lab Color Model)	Percent diff. with RGB Color Model	Percent diff. with HSV Color Model
1-Africa	0.43 (CMW)	0.42 (CMH)	0.60 (F60)	17	18
2-Beach	0.49 (CMH)	0.26 (CMH)	0.45 (F100)	-4	9
3-Building	0.30 (CMH)	0.32 (CMH)	0.45 (F80)	15	13
4-Bus	0.49 (CMH)	0.62 (CMH)	0.55 F(100)	6	7
5-Dinosour	0.99 (CMH)	0.97 (CM)	0.55 F(40)	-45	-42
6-Elephant	0.51 (CM)	0.41 (CM)	0.40 F(100)	-11	-1
7-Flower	0.73 (CMH)	0.60 (CMH)	0.95 F(80)	22	35
8-Horse	0.88 (CMH)	0.43 (CMW)	0.45 F(60)	-45	-43
9-Mountain	0.40 (CMH)	0.38 (CMH)	0.70 (F100)	20	32
10.Food	0.50 (CMW)	0.42 (CMH)	0.75 (F100)	25	33
Average IPR	0.57	0.48	0.59	-	-

where CMW=Color moment extracted the whole image; CMH=Color moment extracted the vertical image; CM=Color moment extracted the horizontal image; F=Frame

Table 3: Comparison of performance of precision for results of 20 image retrievals for all categories using a quantized color histogram (CH) with the proposed method

Category	Ayan et al. (2016)			Proposed Method	Performance Comparison		
	A	B	C	The best performance of Precision	Percent (%) diff. with A	Percent (%) diff. with B	Percent (%) diff. with C
0.571-Africa	0.59	0.67	0.79	0.60 (F60)	1	-7	-19
2-Beach	0.29	0.38	0.30	0.45 (F100)	16	7	15
3-Building	0.36	0.43	0.45	0.45 (F80)	9	2	equal
4-Bus	0.67	0.68	0.63	0.55 (F100)	-12	-13	-8
5-Dinosour	0.84	0.89	0.90	0.55 (F40)	-29	-34	-35
6-Elephant	0.38	0.40	0.36	0.40 (F100)	2	equal	4
7-Flower	0.62	0.56	0.46	0.95 (F80)	33	39	49
8-Horse	0.81	0.85	0.85	0.45 (F60)	-36	-40	-35
9-Mountain	0.33	0.31	0.28	0.70 (F100)	37	39	42
10-Food	0.60	0.63	0.64	0.75 (F100)	15	12	11
% Average IPR					36	5	24

where A=CH (8,2,2), B=CH (8,3,3), C=CH (16,4,4)

Acknowledgment

Compliance with ethical standards

Conflict of interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

Arbelaitz O, Gurrutxaga I, Muguerza J, Pérez JM, and Perona I (2013). An extensive comparative study of cluster validity indices. Pattern Recognition, 46(1): 243-256. https://doi.org/10.1016/j.patcog.2012.07.021

Ayan M, Erdem OA, and Bilge HŞ (2016). Multi-featured content-based image retrieval using color and texture features. In: Dursun M (Ed.), Applied mechanics and materials: 136-143. Volume 850, Trans Tech Publications Ltd., Freienbach, Switzerland. https://doi.org/10.4028/www.scientific.net/AMM.850.136

Baji F and Mocanu M (2017). Connected components objects feature for CBIR. In the 18th International Carpathian Control Conference, IEEE, Sinaia, Romania: 545-550. https://doi.org/10.1109/CarpathianCC.2017.7970460

Bchir O, Ismail MMB, and Aljam H (2018). Region-based image retrieval using relevance feature weights. International Journal of Fuzzy Logic and Intelligent Systems, 18(1): 65-77. https://doi.org/10.5391/IJFIS.2018.18.1.65

Burger W and Burge MJ (2016). Histograms and image statistics. In: Burger W and Burge MJ (Eds.), Digital image processing: 37-56. Springer, London, UK. https://doi.org/10.1007/978-1-4471-6684-9_3

Carson C, Thomas M, Belongie S, Hellerstein JM, and Malik J (1999). Blobworld: A system for region-based image indexing and retrieval. In the International Conference on Advances in Visual Information Systems, Springer, Amsterdam, Netherlands: 509-517. https://doi.org/10.1007/3-540-48762-X_63

Chaimontree S, Atkinson K, and Coenen F (2010). Best clustering configuration metrics: Towards multiagent based clustering. In the International Conference on Advanced Data Mining and Applications, Springer, Chongqing, China: 48-59. https://doi.org/10.1007/978-3-642-17316-5_5

Chan YK, Ho YA, Liu YT, and Chen RC (2008). A ROI image retrieval method based on CVAAO. Image and Vision Computing, 26(11): 1540-1549. https://doi.org/10.1016/j.imavis.2008.04.019

Chaudhuri S (2007). Segmentation and region of interest based image retrieval in low depth of field observations. Image and Vision Computing, 25(11): 1709-1724. https://doi.org/10.1016/j.imavis.2006.12.020

Chen W, Li Q, and Dahal K (2015). ROI image retrieval based on multiple features of mean shift and expectation–maximization. Digital Signal Processing, 40: 117-130. https://doi.org/10.1016/j.dsp.2015.01.003

Eze P, Parampalli U, Evans R, and Liu D (2019). Integrity verification in medical image retrieval systems using spread spectrum steganography. In the International Conference on Multimedia Retrieval, Association for Computing Machinery, Ottawa, Canada: 53-57. https://doi.org/10.1145/3323873.3325020

Hossain S and Islam R (2017). A new approach of content based image retrieval using color and texture features. British Journal of Applied Science and Technology, 21(3): 1-16. https://doi.org/10.9734/BJAST/2017/33326

Jain AK and Dubes C (1988). Algorithms for clustering data. Prentice-Hall, Hoboken, USA.

Kam AH, Ng TT, Kingsbury NG, and Fitzgerald WJ (2000). Content based image retrieval through object extraction and querying. In the Workshop on Content-Based Access of Image and Video Libraries, IEEE Computer Society, Hilton Head, USA: 91-91. https://doi.org/10.1109/IVL.2000.853846

Kannan K, Kanna BR, and Aravindan C (2010). Root mean square filter for noisy images based on hyper graph model. Image and Vision Computing, 28(9): 1329-1338. https://doi.org/10.1016/j.imavis.2010.01.013

Keilwagen J, Grosse I, and Grau J (2014). Area under precision-recall curves for weighted and unweighted data. PLOS ONE, 9(3): e92209. https://doi.org/10.1371/journal.pone.0092209 **PMid:24651729 PMCid:PMC3961324**

Kim S, Park S, and Kim M (2004). Image classification into object/non-object classes. In the International Conference on Image and Video Retrieval, Springer, Dublin, Ireland: 393-400. https://doi.org/10.1007/978-3-540-27814-6_47

Lee YS, Hao SS, Lin SW, and Li SY (2012). Image retrieval by region of interest motif co-occurence matrix. In the International Symposium on Intelligent Signal Processing and Communications Systems, IEEE, Tamsui, Taiwan: 270-274. https://doi.org/10.1109/ISPACS.2012.6473494

Li J and Wang JZ (2003). Automatic linguistic indexing of pictures by a statistical modeling approach. IEEE Transactions on Pattern Analysis and Machine Intelligence, 25(9): 1075-1088. https://doi.org/10.1109/TPAMI.2003.1227984

Liu Y, Zhang D, Lu G, and Ma WY (2007). A survey of content-based image retrieval with high-level semantics. Pattern Recognition, 40(1): 262-282. https://doi.org/10.1016/j.patcog.2006.04.045

Malviya N, Choudhary N, and Jain K (2017). Content based medical image retrieval and clustering based segmentation to diagnose lung cancer. Advances in Computational Sciences and Technology, 10(6): 1577-1594.

Mamat AR, Awang MK, Rawi NA, Awang MI, and Kadir MFA (2016a). Average analysis method in selecting Haralick's texture features on color co-occurrence matrix for texture based image retrieval. International Journal of Multimedia and Ubiquitous Engineering, 11(2): 79-88. https://doi.org/10.14257/ijmue.2016.11.2.10

Mamat AR, Mohamed FS, Mohamed MA, Rawi NM, and Awang MI (2018). Silhouette index for determining optimal k-means clustering on images in different color models. International Journal of Engineering and Technology, 7(2.14): 105-109. https://doi.org/10.14419/ijet.v7i2.14.11464

Mamat AR, Mohamed FS, Rawi NA, Awang MK, Awang M, and Kadir MFA (2016b). Region based image retrieval based on texture features. Journal of Theoretical and Applied Information Technology, 92: 9-19.

Mamat AR, Rawi NA, Awang MI, Kadir MFA, and Rahman MNA (2015). Hybrid method to obtain interest region and non interest region for color based image retrieval. International Journal of Advances in Soft Computing and Its Applications, 7(3): 16-30.

Manning CD, Raghavan P, and Schütze H (2009). An introduction to information retrieval. Cambridge University Press, Cambridge, UK.

Manoharan S and Sathappan S (2013). A novel approach for content based image retrieval using hybrid filter techniques. In the 8th International Conference on Computer Science and Education, IEEE, Colombo, Sri Lanka: 518-524. https://doi.org/10.1109/ICCSE.2013.6553965

Maulik U and Bandyopadhyay S (2002). Performance evaluation of some clustering algorithms and validity indices. IEEE Transactions on Pattern Analysis and Machine Intelligence, 24(12): 1650-1654. https://doi.org/10.1109/TPAMI.2002.1114856

Moghaddam B, Biermann H, and Margaritis D (2001). Regions-of-interest and spatial layout for content-based image retrieval. Multimedia Tools and Applications, 14(2): 201-210. https://doi.org/10.1023/A:1011355417880

Prasad BG, Biswas KK, and Gupta SK (2004). Region-based image retrieval using integrated color, shape, and location index. Computer Vision and Image Understanding, 94(1-3): 193-233. https://doi.org/10.1016/j.cviu.2003.10.016

Raja R, Kumar S, and Mahmood MR (2020). Color object detection based image retrieval using ROI segmentation with multi-feature method. Wireless Personal Communications, 112(1): 169-192. https://doi.org/10.1007/s11277-019-07021-6

Rejito J, Abdullahi AS, Setiana D, and Ruchjana BN (2017). Image indexing using color histogram and k-means clustering for optimization CBIR in image database. Journal of Physics: Conference Series-IOP Publishing, 893: 012055. https://doi.org/10.1088/1742-6596/893/1/012055

Shrivastava N and Tyagi V (2015). A review of ROI image retrieval techniques. In the 3rd International Conference on Frontiers of Intelligent Computing: Theory and Applications, Springer, Bhubaneswar, India: 509-520. https://doi.org/10.1007/978-3-319-12012-6_56

Singh SM and Hemachandran K (2012). Content-based image retrieval using color moment and Gabor texture feature. International Journal of Computer Science Issues, 9(5): 299-309.

Szeliski R (2010). Computer vision: Algorithms and applications. Springer Science and Business Media, Berlin, Germany.

Talib J (2006). Struktur matematik diskret-untuk sains komputer. Penerbit UTM, Skudai, Malaysia.

Tian Q, Wu Y, and Huang TS (2000). Combine user defined region-of-interest and spatial layout for image retrieval. In the International Conference on Image Processing (Cat. No. 00CH37101), IEEE, Vancouver, Canada, 3: 746-749. https://doi.org/10.1109/ICIP.2000.899562

Toriwaki J and Yoshida H (2009). Fundamentals of three-dimensional digital image processing. Springer Science and Business Media, Berlin, Germany. https://doi.org/10.1007/978-1-84800-172-5

Vimina ER and Jacob KP (2013). A sub-block based image retrieval using modified integrated region matching. International Journal of Computer Science Issues, 10(1): 686-692. https://doi.org/10.12720/joig.1.1.7-11

Vu K, Hua KA, and Tavanapong W (2003). Image retrieval based on regions of interest. IEEE Transactions on Knowledge and Data Engineering, 15(4): 1045-1049. https://doi.org/10.1109/TKDE.2003.1209021

Wang Z, Liu G, and Yang Y (2013). A new ROI based image retrieval system using an auxiliary Gaussian weighting scheme. Multimedia Tools and Applications, 67(3): 549-569. https://doi.org/10.1007/s11042-012-1059-3

Yang L, Geng B, Cai Y, Hanjalic A, and Hua XS (2011). Object retrieval using visual query context. IEEE Transactions on Multimedia, 13(6): 1295-1307. https://doi.org/10.1109/TMM.2011.2162399

Yue J, Li Z, Liu L, and Fu Z (2011). Content-based image retrieval using color and texture fused features. Mathematical and Computer Modelling, 54(3-4): 1121-1127. https://doi.org/10.1016/j.mcm.2010.11.044

Zambre DS and Patil SP (2013). Retrieving content based images with query point technique based on k-mean clustering. International Journal of Advancements in Research and Technology, 2(4): 357-360.

Zhang D and Lu G (2002). Generic Fourier descriptor for shape-based image retrieval. In the International Conference on Multimedia and Expo, IEEE, Lausanne, Switzerland, 1: 425-428. https://doi.org/10.1109/ICME.2002.1035809

Zhou Q, Ma L, Celenk M, and Chelberg D (2005). Content-based image retrieval based on ROI detection and relevance feedback. Multimedia Tools and Applications, 27(2): 251-281. https://doi.org/10.1007/s11042-005-2577-z