Region of interest-based image retrieval techniques: a review

Mardhiyah Md Jan¹, Nasharuddin Zainal², Shahrizan Jamaludin³

¹²Department of Electrical, Electronic & Systems Engineering, Universiti Kebangsaan Malaysia, Malaysia
³Department of Maritime Technology, Universiti Malaysia Terengganu, Malaysia

ABSTRACT

This paper presents a review of the region of interest-based (ROI) image retrieval techniques. In this study, the techniques, the performance evaluation parameters, and databases used in image retrieval process are being reviewed. A part of an image that is considered important or a selected certain area of the image is what defines a region of interest. Retrieval performance in large databases can be improved with the application of content-based image retrieval systems which deals with the extraction of global and region features of images. The capability of reflecting users’ specific interests with greater accuracy has shown to be more effective when using region-based features compared to global features. Segmentation, feature extraction, indexing, and retrieval of an image are the tasks required in retrieving images that contain similar regions as specified in a query. The idea of the region of interest-based image retrieval concepts is presented in this paper and it is expected to accommodate researchers that are working in the region-based image retrieval system field. This paper reviews the work of image retrieval researchers in the span of twenty years. The main goal of this paper is to provide a comprehensive reference source for scholars involved in image retrieval based on ROI.

This is an open access article under the CC BY-SA license.

1. INTRODUCTION

The rapid expansion of the internet and fast advancement in colour imaging technologies have made digital colour images more readily obtainable to professional and amateur users [1]. The medical field, surveillance system, and digital forensics are among the areas that have been using a vast amount of multimedia data in the form of audio, video, and images in this advance development of internet and multimedia technologies. This situation leads to the need for developing a system that is able to store and efficiently retrieve these digital data [2].

The main goal of an image retrieval system is to search and retrieve images from a variety and large databases in minimal time with high accuracy [3]. Two of the techniques that are generally used to search and retrieve images in the database are the Text-based Image Retrieval (TBIR) and Content-based Image Retrieval (CBIR). TBIR works by using retrieval keys which include classification codes, keywords or subject headings in retrieving images. TBIR is considered a non-standard technique of retrieving images due to the inconsistency of interpretation of keywords used by users.
However, TBIR is the most common retrieval technique that is being used, where the search is based on the explanation of the images. Normally a TBIR works by running a database search for similarity text surrounding the images as given by the query string of any keyword used by users. At present, Google Images is the most frequently used TBIR system. The string matching requires less computation time making the TBIR system works fast. Nevertheless, there is also a disadvantage of TBIR where it is at times difficult to describe the entire graphic content of images in word form. This factor may lead to the TBIR system to produce irrelevant results when querying. Some annotations of the images may also be incorrect and this will consume a lot of time for the TBIR system to yield the desired output. TBIR can be successfully defined as a document retrieval problem [4].

CBIR systems were developed to overcome the limitations met by the TBIR systems. CBIR is one of the instances of information retrieval that applies computer vision techniques to solve searching and managing large image databases related problems [5]. A CBIR uses graphic contents of a certain image defined by colour texture, shape, and spatial location that are identified as the low-level features to represent the image in the database. When a desired image is being used as an input in the system, it will retrieve a similar image matched to the example image provided. By querying using CBIR systems it will eliminate the requirement for expressing the graphic content images in words form that resembles the human perception of visual data. Among the CBIR systems that are commonly available are QBIC [6], Photobook [7], VisualSeek [8], Netra [9] and SIMPLicity [10].

There have been extensive studies done on CBIR and the progress has been discussed comprehensively in [11-13]. An efficient shape descriptor is essential to identify the features of an image content, including its shape, colour and texture [14-15]. Region-based Image Retrieval (RBIR) system is recognized as one of the categories of the CBIR system. Image segmentation and defining measurement criteria constitute the ROI-based image retrieval system [16]. This system works by producing the representation of the image in the database by utilizing the features extracted from the region or part of the image.

RBIR systems can be categorized into System Designated Region-of-Interest (SDR) [17] and User Designated Region-of-Interest (UDR) [18] approaches. When querying to the database using the SDR approach, dividing the image into significant regions and designating every region as ROI is done by the system automatically. Meanwhile, by using the UDR method, the selection of ROI in the image for query formulation is done manually by the user.

One of the factors to determine the success of an SDR method is the accuracy of the segmentation technique used in dividing an image into many regions. However, unpredicted noise in the output may result in the reduction of retrieval accuracy thus making image segmentation not always reliable. Inquest to identify either the boundaries of the regions of the objects in the image are the two diverse yet complementary perspectives that are usually being used in image segmentation [19]. Furthermore, many existing segmentation techniques fail to extract objects of interest despite their ability to accurately identify specific regions from images.

The SDR method has some limitations in reflecting the user's goal in the process of retrieval due to these reasons. It is complicated to decide beforehand the part of the image that will be chosen when the user opts for the ROI using the UDR approach manually. Extraction of the images' feature values and matching them with the ROI for retrieval is done by dividing the image into a smaller number of blocks is the solution to this problem. Correct selection of blocks overlaying the ROI is a crucial decision since the UDR method is able to have a variety of sizes and can have multiple blocks in order to satisfy the user query precisely [20].

ROI placement also plays a significant role in obtaining an effective ROI image retrieval [21]. Fixed location matching will be the result in the case of blocks having the exact placement as the ROI. However, this method could not retrieve the same images when similar regions to ROI lie in other parts of the images. For instance, the system fails to produce the exact images containing an elephant that is situated on the top-left corner of an image when initially the user who queries for an elephant in the bottom-right corner of the block image is being used. Opting for all blocks matching strategy might solve this problem, but using this approach will lead to the time complexity that needed to be compensated as it increases the computational complexity and time as the blocks increases in dimension layout. This paper presents the review of region of interest-based image retrieval with various of retrieval techniques used and the findings obtained.

2. IMAGE PROCESSING METHODS

Image processing generally relies on the specific feature of an image and matching it to any specific object. This method is performed to extract useful information from the images. Digital image processing focuses on improving pictorial information for human interpretation and processing of image data for transmission and storage. Recognition of an object is a very significant challenge in image processing.
Since every object is a part of an image, thus, matching ROI can also be considered as an object of a particular image. There are several operations involved in image processing methods such as representation, segmentation, clustering, feature extraction, and matching. These operations will further be discussed below.

2.1. Image representation
There are a lot of images available in the digital world today, however, the actual representation of an image is denoted as a collection of discrete picture element called as the pixels. Pixel values are most often grey levels in the range of 0 until 255 (black-white). Binary is the image representation format usually used and later will be translated into different image formats such as jpg, bitmap, and png [22-23]. These image representations rely on the pixel size and these images are represented by pixels similar to building blocks. In order to reduce the complexity and computational time of image processing, a high-quality image is required [24].

2.2. Image segmentation
The definition of image segmentation in computer vision is given as the process of segregating a digital image into many segments. Image segmentation partitions the image pixels into groups that strongly correlate with the objects in an image. The aim of segmentation is to simplify the image into a representation that is easier to analyse and more meaningful [19, 25]. In the image retrieval field, the segmentation process is still considered an open problem in computer vision despite some of the systems that can deliberately be very efficient. By only segmenting the region required by the user instead of having to do the segmentation process for the whole image, the ROI concept that is used can reduce the computation time since the segmented region is smaller compared to the entire image.

2.3. Region clustering
In terms of features present, there might be features that are quite similar in many regions that belong to a different set of images [26]. A training data and a codebook size are produced by getting the average of efficacy and accuracy of the retrieval based on the features that are obtained from the regions of the whole database images. All the data and details are computed then stored as a codeword. The cluster is recognized and will be registered in the codebook and the matching index will be stored for each region of an image in the database used while removal will be done to the original feature of that particular region. The nearest entry in the codebook is located and the matching index will be replacing its feature for a new image region. Region clustering rely mainly on the assumption that the pixels of one region is similar with the neighbouring pixels.

2.4. Feature extraction
Feature extraction is important since it is the process of obtaining global image features such as colour or local descriptors like texture and shape of a particular image. In this phase, an identity is given for each character which is represented by a feature vector [27]. When being compared to dimensional histograms, these representations are considered more efficient in terms of retrievals. The most prominent features of the image only will be extracted after feature extraction analysis has been done by the system [28]. One of the key attributes of an image region that has been segmented is the shape of the region and it is a crucial role in representing the effectiveness of a robust image retrieval system [29]. The aim of feature extraction is to extract some sets of features in order to minimize the image retrieval rate.

2.5. The general framework of RBIR
User-dependent retrieval and system-dependent retrieval are two of the most commonly used methods that exist in region-based image retrieval. In the UDR approach, features of the blocks in the layout that the user selects will be the region's features that the system will extract. Meanwhile, by using the SDR approach, regions or objects of the image will be the potential ROIs based on the segmentation algorithm. The obtained features will then be utilized for matching purposes to produce results of the image retrieval and producing database images of the region selection [25]. The RBIR works by filtering out the candidate image and calculating the similarity with the query using the image database which will lead to significant improvement in retrieval speed [30].

3. RELATED WORKS
Most of the literature on ROI-based image retrieval considers a variety of issues that can be solved. Many authors have contributed their works in the ROI area and some of them are listed below:
Chen et al. [31], discussed a new approach to index image in content-based image retrieval by including Kohonen’s Self Organizing Map (SOM), one of the neural network model available. They stated that the significant contribution of this approach is that it is interactive, where the user is able to choose an ROI desired from the sample images before the system focuses on the query matching result of the corresponding regional colour features, thus producing the region that contains the same regions as required by the user. The capability of the system to partition every image into many standardize regions for indexing and representing the image adaptively is possible using the SOM algorithm. Figure 1 depicts the output of the system.

The users are the more suitable person to specify the content of the query and not the computer itself, thus Moghaddam et al. [32] presented another alternative image retrieval system based on this opinion. The users are able to choose multiple ROI and can specify the relevance of their spatial layout in the retrieval process by using this proposed system. The similarity bounds on histogram distances for reducing the database search is also derived in this discussion. An example of multiple ROI queries using an MRI database can be seen in Figure 2.

Figure 1. An example of retrieved images containing flowers with yellow colour from the database using [31]
Figure 2. An example of an MRI imagery database using multiple ROI query using the method by [32]

Another technique is the region-based image retrieval system using a high-level semantic learning idea that has been suggested by [33]. The significant feature of this system is its ability to process both query by ROI and query by keyword. The image will be partitioned into several regions and the system will extract the regions with low-level features. High-level features are then obtained using the suggested decision tree-based learning algorithm named decision tree-based image semantic learning algorithm, DT-ST. The examples of query and dominant region using the proposed system are portrayed in Figure 3.

Combining the text-based retrieval system using a unique region-based inverted file indexing method with the semantic image retrieval model is the proposed idea discussed by [34]. By using this method, images were interpreted into textual documents which are then being indexed and retrieved, similar to the conventional text-based query process. Two examples of the output of this proposed retrieval system are shown in Figure 4.

Figure 3. Dominant region query image examples obtained from [33]
4. SUMMARIZED LITERATURE REVIEW

The properties used in 15 articles related to region-based Image Retrieval systems reviews in this paper are listed in Table 1. The properties are classified into the concept, performance evaluation parameter, the database used, and findings by concerned authors.

<table>
<thead>
<tr>
<th>Authors/Year</th>
<th>Technique</th>
<th>Performance Evaluation Parameter</th>
<th>Databases</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>T. Chen, L.-H. Chen, K.-K. Ma [31] (1999)</td>
<td>Region-of-interest (ROI) image retrieval based on colour image indexing</td>
<td>Precision of image retrieval</td>
<td>500 various contents images</td>
<td>High precision in results of image retrieval. Rotated and scaled versions of images can also be detected by using the proposed method.</td>
</tr>
<tr>
<td>Qi Tian, Ying Wu, Thomas S. Huang [35] (2000)</td>
<td>Combining user-defined region of interest (ROI) and spatial layout Content-based Image Retrieval</td>
<td>Accuracy of image retrieval</td>
<td>COREL database images</td>
<td>More capable of capturing image details compared to the global approach.</td>
</tr>
<tr>
<td>Baback Moghaddam, Henning Biermann, Dimitris Margaritis [32] (2001)</td>
<td>Multiple region-of-interest (ROI) and spatial layout for Content-based Image Retrieval</td>
<td>Accuracy and speed of image retrieval</td>
<td>VisTex and GIS orthophoto Imagery database images</td>
<td>The proposed method can reduce computational time and speed-up the retrieving possibility based on partial spatial configuration.</td>
</tr>
<tr>
<td>Feng Jing, Ming Jing Li, Lei Zhang, Hong-Jiang Zhang, Bo Zhang [36] (2003)</td>
<td>Learning in region-based image retrieval</td>
<td>Accuracy of image retrieval</td>
<td>10 000 COREL database images</td>
<td>Shows the effectiveness of retrieval after comparing results using the learning schemes for region-based image retrieval with database images.</td>
</tr>
<tr>
<td>Khanh Vu, Kien A. Hua, Wallapak Tavanapong [37] (2003)</td>
<td>Region-of-interest (ROI) queries based on sampling matching</td>
<td>Robustness and speed of image retrieval</td>
<td>Art Explosion commercial database images</td>
<td>The user-defined region query can overcome irrelevant region retrieved results. Fast retrieval when the indexing technique is used.</td>
</tr>
</tbody>
</table>
5. CONCLUSION AND FUTURE WORKS

This paper categorizes the various techniques and examples used in region of interest-base image retrieval. A collection of papers was studied and numerous concepts and methods are classified in the given table. Several results of the proposed methods by authors are being presented in the illustrated figures. In the image retrieval field, the development of inexpensive, rapid and robust retrieval systems are among the important factors to be considered. Hence, an immense range of applications can gain benefits from these image retrieval technologies. There are countless field that uses these RBIR techniques applications such as medical imaging, archaeology, zoology, and criminal investigation. Overall, the objective of this study is to provide comprehensive reference for researchers in similar research area has been achieved. In conclusion, there is a lot of research yet to be done to improve the performance used in the current technology of RBIR and to expend the application in many more fields.

Table 1. Summarized literature survey of region-based image retrieval techniques

<table>
<thead>
<tr>
<th>Authors/Year</th>
<th>Technique</th>
<th>Performance Evaluation Parameter</th>
<th>Databases</th>
<th>Findings</th>
</tr>
</thead>
<tbody>
<tr>
<td>B. G. Prasad, K. K. Biswas, S. K. Gupta [38] (2004)</td>
<td>Region-based Image Retrieval using integrated colour, shape and location index</td>
<td>Efficacy of image retrieval</td>
<td>MPEG-7 database images</td>
<td>The proposed method is robust towards scaling and rotation invariance</td>
</tr>
<tr>
<td>Feng Jing, Ming Jing Li, Hong-Jiang Zhang, Bo Zhang [39] (2004)</td>
<td>Region-based image retrieval using relevance feedback</td>
<td>Precision of image retrieval</td>
<td>10 000 COREL Stock Photo Library</td>
<td>Integrating the relevance feedback algorithm into a region-based image retrieval results in a more precise output of the image retrieved.</td>
</tr>
<tr>
<td>Byoung Chul Ko, Hyeon Byun [40] (2005)</td>
<td>Semantic image segmentation using adaptive circular filters based on Bayes' theorem and texture distribution</td>
<td>Region precision of image retrieval</td>
<td>COREL photo CD images</td>
<td>The proposed method able to accommodate region scaling, rotation, and translation.</td>
</tr>
<tr>
<td>Qiang Zhou, Limin Ma, Mehmet Celenk, David Chelberg [41] (2005)</td>
<td>Content-based image retrieval based on region-of-interest (ROI) detection and relevance feedback</td>
<td>Accuracy of image retrieval</td>
<td>2500 natural colour images</td>
<td>The proposed method shows promising results compared to the global feature-based image description.</td>
</tr>
<tr>
<td>Jing Zhang, Choong-Woong Yoo, Seok-Wun Ha [42] (2007)</td>
<td>Region-of-interest (ROI) based image retrieval by combining colour and texture feature</td>
<td>Precision and speed of image retrieval</td>
<td>300 nature images (<a href="http://www.cs.washington.edu/research/imag">http://www.cs.washington.edu/research/imag</a> edatabase/)</td>
<td>The proposed method improves in accuracy and speed of image retrieval.</td>
</tr>
<tr>
<td>Yung-Kuan Chan, Yu-An Ho, Yi-Tung Liu, Rung-Ching Chen [43] (2008)</td>
<td>Colour Variances among Adjacent Objects (CVAAO)</td>
<td>Precision of image retrieval</td>
<td>Full-colour images</td>
<td>CVAOO feature can describe pixel colour and colour complexity of an image. Able to precisely sever the target region image. The method is insensitive towards scale, shift, rotation and distortion variance.</td>
</tr>
<tr>
<td>Dengsheng Zhang, Md Monirul Islam, Guojuan Lu, Jin Hou [34] (2009)</td>
<td>Semantic image retrieval using region base inverted file</td>
<td>Accuracy of image retrieval</td>
<td>7600 COREL database images</td>
<td>The proposed method presents higher retrieval performance compared to traditional content-based image retrieval.</td>
</tr>
<tr>
<td>Chaobing Huang, Qian Liu, Shengsheng Yu [44] (2011)</td>
<td>Region-of-interest (ROI) extraction based on visual saliency</td>
<td>Accuracy of ROI extraction</td>
<td>200 SIMPLcity database images</td>
<td>Discrete Moment Transform (DMT)-based saliency can determine large ROI.</td>
</tr>
<tr>
<td>Zhe Wang, Guizhong Liu, Yang Yang [45] (2012)</td>
<td>Region-of-interest (ROI)-based image retrieval system using an auxiliary Gaussian weighting scheme (AGW)</td>
<td>Precision of retrieval system</td>
<td>100 000 Google Images</td>
<td>High precision retrieval performance However, the proposed method shows inaccuracies towards scaling and rotating image invariances.</td>
</tr>
</tbody>
</table>

ACKNOWLEDGEMENTS

This research activity had been conducted in the Vision@UKM laboratory and is supported by the Ministry of Education (MOE) Grant fund: FRGS/1/2018/TK04/UKM/02/16.

REFERENCES


**BIOGRAPHIES OF AUTHORS**

Mardhiyah Md Jan received her BEng in Electronics Computer & Information Engineering from International Islamic University Malaysia (HUM), Malaysia, in 2013 and MEng in Communication and Computer Engineering from Universiti Kebangsaan Malaysia (UKM), Malaysia, in 2015. Currently she is pursuing her PhD studies at Department of Electrical, Electronics & Systems Engineering, Faculty of Engineering & Built Environment, Universiti Kebangsaan Malaysia, Bangi, Malaysia. Her main research interests include computer engineering, image processing and projection mapping.

Dr. Nasharuddin Zainal received his BEng in Computer Science from Tokyo Institute of Technology (Japan), in 1998, MEng in Communication and Computer Engineering from Universiti Kebangsaan Malaysia (UKM), Malaysia, in 2003, and DrEng degree from Tokyo Institute of Technology (Japan) in 2010. He is currently an Associate Professor at the Department of Electrical, Electronics & Systems Engineering, Faculty of Engineering & Built Environment, Universiti Kebangsaan Malaysia, Bangi, Malaysia. His main research interests include computer engineering, image and video processing and pattern recognition.

Dr. Shahrizan Jamaludin received his BEng in Electronic (Computer) Engineering from Universiti Teknikal Malaysia Melaka (UTeM) and both his MEng in Communication and Computer Engineering and PhD in Electrical, Electronic and Systems Engineering from Universiti Kebangsaan Malaysia (UKM). He is a graduate engineer certified by the Board of Engineers Malaysia and a graduate technologist certified by the Malaysia Board of Technologists. Currently, he is a senior lecturer at the Department of Maritime Technology, Faculty of Ocean Engineering Technology and Informatics, Universiti Malaysia Terengganu (UMT). His research interests include Image processing, computer vision, signal processing, pattern recognition, embedded system, computer engineering, software engineering, control system, automation, instrumentation, robotic, maritime technology and occupational safety and health.