Region-Based Image Retrieval Using Relevance Feature Weights

Ouiem Bchir, Mohamed Maher Ben Ismail, and Hadeel Aljam

Department of Computer Science, College of Computer and Information Sciences, King Saud University, Riyadh, Saudi Arabia

Ijfis

Abstract

We propose a new region-based CBIR (content-based image retrieval) system. One of the main objectives of our work is to reduce the semantic gap between the visual characteristics of the query and the high level semantic sought by the user. This is achieved by allowing the user to select specific regions and expressing his interest in a more accurate way. Moreover, the proposed approach overcomes the challenge of choosing suitable features to describe the image content. More specifically, relevance weights are automatically associated with each visual feature in order to better represent the visual content of the images. To evaluate these objectives, we compare the obtained results with those obtained using traditional CBIR systems.

Keywords: Region-based retrieval, Content-based retrieval, Feature weighting

1. Introduction

Nowadays, the increasing use of multimedia technology, electronic devices equipped with digital cameras, and the billions of photos uploaded on social media lead to the continued exponential growth of digital image databases. This exponential growth has triggered the need for efficient retrieval techniques aiming at finding images of interest. Text-based image retrieval (TBIR) is the earliest and most common retrieval technique. It is based on keywords or description surrounding images, and that are used to describe image content. When the text query is submitted, TBIR system retrieves the images that are annotated with similar keywords. Popular search engines, such as Google, have first used TBIR to retrieve images. TBIR exploits string matching and text mining techniques to allow fast and reasonable image retrieval. TBIR approach relies on the assumption that all images are annotated. However, this assumption is not held for large digital image databases where users upload their own files without adding specific names or labels. A natural solution to overcome this disadvantage is manual labeling of the images. However, this solution represents a tedious work for small image collection, and impractical for large databases. Another drawback of TBIR approach consists in the semantic gap between the user text query, and the visual properties of images. For instance if the user submits the query “apple” then he may be looking for the company Apple [1] or the fruit apple. The alternative to overcome this limitation consists in bridging the gap between the semantic of text query and the visual content of images by allowing the user to formulate the visual properties of the images he is looking for.

Content-based image retrieval (CBIR) emerged as a novel approach to efficiently retrieve

Received: Jan. 29, 2018
 Revised : Mar. 17, 2018
 Accepted: Mar. 19, 2018

Correspondence to: Mohamed Maher Ben Ismail
 (maher.benismail@gmail.com)
 ©The Korean Institute of Intelligent Systems

©This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

relevant images. CBIR techniques tends to retrieve the images that are visually similar to a given query image. For typical CBIR system, the user provides the system with an image as a query, and seeks relevant images. This retrieval process encodes and indexes the images using their visual features, and the returned images are selected based on the similarity of their visual features to those representing the query. Thus, the retrieval performance depends on the choice of the visual low-level features. In other words, the selection of the low-level feature represents the keystone of typical CBIR systems. CBIR techniques find applications in various research fields such as information retrieval, computer vision and pattern recognition. Several CBIR systems have been developed, and represent promising solutions for searching and retrieving images. Some of the earliest CBIR systems are query by image and video content QBIC system [2], VisualSEEK [3] and SIMPLicity [4], etc.

Despite the publication of various visual features, to report the performance of the visual features used to encode images in CBIR systems, and the significant advances in the image retrieval field. These systems show some limitations due to the extraction of the visual features from the whole image, which yields a semantic gap between the image's high-level meaning and the extracted low-level features. These visual features insufficiently describe some relevant objects that the user is interested in. In fact, images, in general, contain irrelevant areas like the background [5]. In other words, these systems do not succeed to interpret the semantic meaning of the query submitted by the user. In particular, the user is usually interested in a specific region/object in the query image rather than the whole image.

Another limitation of existing CBIR systems is their sensitivity to the type of visual features extracted from the images. In fact, the visual feature relevances depend on the image content. Thus, the more distinctive the visual features are, the more accurate the retrieval result gets. However, existing CBIR systems do not differentiate between low-level features, and when they aggregate several features, a curse of dimensionality may raise as a problem. Actually, allowing the user to select a specific region would narrow this semantic gap, and let the user express his interest in a more accurate way.

In this paper, we propose a novel Region-based CBIR system. More specifically, based on the region selected by the user, we intend to learn relevance weights for each considered visual feature. In particular, we propose a novel approach to estimate the similarity between the region of interest in the query im-

age, and the images in the database. Notice that the approach associates relevance weights to the different low-level features in an unsupervised manner in order to match the user interest. The proposed technique would increase the image retrieval performance and reduce the semantic gap between the visual properties of the query and the semantic meant by the user.

The rest of the paper is organized as follows: Section 2 represents a background about image retrieval approaches and feature extraction methods. Section 3 reviews existing CBIR systems and region-based image retrieval (RBIR) systems. In Section 4, the proposed RBIR approach using relevance feature weights is described. Section 5 outlines the experiments conducted to assess the performance of the proposed approach. Finally, in Section 6 we conclude this work and outline the future works.

2. Background

Typically, CBIR systems rely on image content to search and retrieve similar images. CBIR system consists of two main components. The first one represents the offline phase. It is also called data insertion component. It includes the preprocessing task to enhance the image quality for specific retrieval purposes. The extracted features are then used to represent the database image. The second component of CBIR system is the online phase that is also called query-processing phase. In this step, the user submits his query image through a user interface. Then, the system extracts the visual feature from the query image, and computes the similarity between the query visual features and the database images in the considered feature space. Finally, the system displays the retrieved images.

Visual feature extraction is one of the essential components of any CBIR system. In other words, the accuracy of such CBIR system depends on the power of these visual features to be loyal to the image content. In fact, the image content is encoded into a numerical vector. The resulting vectors represent the image describe image content such as color and texture features. Thus, the feature extraction remains a challenging task. The more representative the feature vectors are, the retrieval results get better. The image visual features can be extracted as global from the whole image or local from regions or objects obtained by image segmentation [6].

3. Related Works

3.1 Content-Based Image Retrieval

CBIR techniques find application in various fields such as information retrieval [7], computer vision and pattern recognition [8]. Several CBIR systems have been developed, and represent promising solution for searching and retrieving images. However, these systems have some limitation due to the extraction of the visual features from the whole image, which yields a semantic gap between the image high-level meaning and the extracted low-level features. Some existing CBIR try to reduce this semantic gap using machine learning techniques [9], user interaction [3] or image segmentation [10]. In the following, we provide a review of existing CBIR techniques.

The authors in [2] developed one of the earliest CBIR system. Namely, they introduced QBIC that allows query by image and video content. QBIC offers large image and video databases that can be mined using sample images, drawings, user-sketches or color and texture patterns. QBIC consists of two main components: database population and database query. The database population aims at processing images and videos in order to extract features such as color, texture, shape, and target motion. Then, the resulting feature vectors are saved in databases. For the database query component, the user formulates his query graphically and the extracted visual features are sent to a matching engine in order to find similar images or videos in the database.

In [3], VisualSEEK system has been proposed as novel CBIR solution allowing image database query using color/spatial properties. It uses image regions of interest and their visual properties to determine relevant database images. This system integrates two queries, which are content-based and spatial based to provide accurate encoding the color and spatial information seek by the user. The system lets the users to look for images by providing the spatial properties and the visual features the user is interested in. Also, it offers two types of query: a single region-based query and a multiple regions-based query. For single region-based query, the region location, along with the area and the spatial extent has to be provided by the user. Moreover, he can assign relative weight to each of these attributes of the region of interested. On the other hand, multiple regions based query joins individual regions queries, and the retrieval results are then obtained by combining the individual results obtained using individual region queries. VisualSEEK system offers image retrieval from a collection of 12,000 digital images. Also, 500 synthetic images by are provided to express the color/space based query. The reported experiments demonstrate that the

color/spatial query outperforms the histogram based retrieval.

The researchers in [11] seek to increase the accuracy of CBIR performance by using gradient projection to represent gray level of the images. The gradient is computed at each pixel in the image and the pixels with a magnitude larger than the thresholds are assigned a value of one. The gradient projection is computed in different directions. Namely, these directions are vertical, horizontal and diagonal. The vectors obtained using the gradient projections are then used to compute the similarity between the two images using the Euclidean distance. The experiments were performed on three databases; the first one contains 1,000 photographic images. The second database encloses 82 camera pictures, and the third one includes a set of object photos. The reported accuracy of testing result was 100% for the three databases. However, the main limitation of this method is that it is restricted to the gray images.

In [12], the authors proposed a CBIR system which relies on object extraction via image segmentation. The query strategy consists of totally automated and semi-automated modes. The experiments performed on a small image database showed that the retrieval performance for “Sunset” and “Zebra” categories is particularly high, whereas “Tiger” and “Autumn scene” classes yield reasonable performance. The authors compared the results of the system without pre-segmentation (global histogram) stage for some image categories such as “Bear”, “Leopard” and “Sunset”. The pre-segmentation approach yielded better result.

The researchers in [9] proposed a cluster-based retrieval approach which uses the unsupervised learning algorithm CLUE [9]. The main contribution in this approach is that it allows the retrieval of image groups rather than a series of ordered images. It takes into consideration the similarity of the target images. The query image and the closest target images are chosen based on a predefined distance metric. It adopts a graph theoretic algorithm to produce groups and the nearest neighbor method (NNM) to select neighbor target images. The experiment proves that CLUE can find semantically more relevant clues than existing CBIR system using the same distance metric. One of the limitations of this approach is the way to define a representative image for a group. It fails sometimes to find semantically relevant representative image.

The authors in [10] proposed a CBIR approach that is based on statistical color histogram feature. This approach gives the same importance to the three components of color space the retrieval process. The authors constructed one probability histogram per color channel. Then, they generated set of relevant bins, and computed several statistical values. Namely, they used

the skewness, the standard deviation and the kurtosis which represent a feature vector for each image. The database used for experiment includes 1,000 images grouped in 10 classes. Each one encloses 100 images. The performance of this CBIR system was measured using precision and recall. It reached a precision of 100% for categories such as dinosaur and “horse”. On the other hand, for “mountain” and “building” classes, it yielded poor performance due to the intra category variance of the color of these two classes.

3.2 Region-/Object-based Image Retrieval

Region-/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. Also it can reduce the semantic gap between the visual properties of the query and the high level understanding of the user. RBIR/OBIR systems segment the images into regions such as in [13–17] or into blocks [18–21] as first step. Then the visual feature locally extracted from each segmented region or block in image. Region-based feature become more effective than global feature and it can reflect user interest. The region-of-interest (ROI) in an image can be specified manually by the user [18–21] or using automatic segmentation technique in [13–16]. Some CBIR/RBIR systems assign weight manually to each considered visual feature [3, 12, 21] to differentiate between their importance. However, in practice this may not reflect the semantic meaning of the query submitted by the user. In the following, we review the state of the art of existing RBIR/OBIR techniques.

The authors in [13] proposed a new method for region-based image similarity calculation. First, a combination of features is considered to calculate the similarity between two regions. Features such as color, texture and shape features are extracted before combination. After that, a weight product method is adopted to obtain the similarity between the two regions. Finally, the average similarity between all regions is considered as the similarity between the two images.

The researchers in [14] proposed a region based image retrieval system which aims at learning high level semantic that reinforces the keyword based query, and the ROI based query. The system starts by segmenting an image into a set of regions. Next, low level visual features (color and texture) are extracted from the regions in the HSV space. Then, high-level concept is obtained from these features using a learning algorithm based on decision tree named DT-ST.

In [15], the authors introduced two context object retrieval (COR) models to overcome object based retrieval challenges. They used SIFT [22] as visual feature for their system. Visual words are obtained using k -means [23] algorithm along with quantized SIFT visual features. A rectangular bounding box is placed on the query image in order to specify the query object.

The researchers in [16] introduced RBIR system which uses the discrete wavelet transform (DWT) [24, 25] in the HSV color space and k -means learning algorithm [26] for image segmentation. The image regions are denoted by a set of visual features and the similarity between regions is estimated using the Bhattacharyya measure.

Recently, the authors in [17] developed new method of OBIR, which relies on a multi-graph multi-instance learning method. The graphs have been created at two levels: the region and the image levels. They also developed a unified optimization framework to exploit the available label information and the graph structure in a comprehensive way. The consistency of annotated images, their assigned labels along with their segmented regions are used and encoded using the mutual restrictions for designing the cost items to build effective relationship between them.

In [18], the authors proposed similarity measure based on relevant regions only. They determine ROI using SamMatch [27] framework. The proposed similarity model depend on the distance between the color of the blocks of two sub-images and the weight factors which are set to indicate the significance of a match at certain block. Also, they proposed an indexing technique by using clustering and an R^* -tree [26] as indexing structure.

Recently, a multilevel indexing structure (MIS) for OBIR has been introduced in [19]. The proposed system is OBIR system that reduces the object search space using a three level tree structure. In addition, in order to speed up the object scanning in the database level, a clustering process is performed in the dataset domain to summarize the dataset. For the retrieval phase, the user selects an object in a given image as query object and the corresponding images which the sub-images correspond the query object are retrieved from the dataset as candidate results.

Similarly, the authors in [20] present an algorithm for ROI retrieval. Their algorithm uses a new image segmentation algorithm which relies on mutation information approach to split images into several target areas. The authors define a new similar measure criterion for the retrieval step where the user selects a ROI, and the ROI-based retrieval finds candidate images

Lately, the authors in [21] proposed a RBIR system with

Figure 1. System architecture.

no automatic segmentation. They used two features: texture and color features in irregular ROIs. The color features were derived using k -means [23] clustering algorithm, and the texture feature consists in Haralick's texture visual feature [28]. In addition, they used an indexing technique that uses a binary code for each region in the image proposed in [28] which allows efficient matching of the queried region. The system measures the similarity using the Euclidian distance between the ROI and the candidate regions in the database images.

As we can notice, the existing CBIR or RBIR/OBIR do not take into account automatic relevance weighting for the extracted features, which effects the retrieval process. In fact, the more distinctive the visual features is, the more accurate retrieval result we get. In this work, we seek to increase the performance of image retrieval and reduce the semantic gap by learning relevance weights for the different low-level features in an unsupervised manner in order to match the user interest.

4. Proposed Approach

As stated in the related work section, the main limitation of the existing CBIR/RBIR systems consists in the fact that they do not discriminate between the visual features in an unsupervised manner to better capture the user semantic. In fact, whenever the visual features get more distinctive, the retrieval result would get more accurate. Therefore, we propose to design and implement a region-based CBIR system which learns a relevance weight for each visual feature to better represent the visual content of the image. Therefore, we propose a novel approach to estimate the similarity between the ROI in the query image, and the images in the database. Our approach associates relevance weights to the different low-level features in an unsupervised manner in order to match the user interest. Using the proposed technique, we aim at increasing the retrieval performance and reducing the semantic gap between the visual characteristics of the query and its semantic meaning.

Figure 1 gives an overview of the proposed system. As it can be seen, it consists of two main phases. Namely, the offline phase shown in Figure 1(a), and the online phase shown in Figure 1(b). During the offline phase, each database image is segmented into regions. Then, the visual features are then extracted from each image region. These visual features are used to represent each region in the feature space. In the online phase, the user submits a query region, and the system automatically extracts the visual features. Then, the system learns the relevance weights for each visual feature in an unsupervised

manner as outlined in the next section. Next, the similarity between the query region submitted by the user and image regions in the database is computed. Finally, the system retrieves top M images corresponding to the highest similarity values.

4.1 Unsupervised Weight Learning

In this section, we present how the proposed system learns the relevance weights in an unsupervised manner to estimate the similarity between the ROI in the query image, and the images in the database.

Let i represents the region selected by the user from the query image, and j represents the image in the database that has been segmented into K regions $\{R_{d1}, \dots, R_{dK}\}$. Let $f_i = \{f_i^1, \dots, f_i^s\}$ be the set of features extracted from query region. Similarly, let $f_{jk} = \{f_{jk}^1, \dots, f_{jk}^s\}$ be the set of features extracted from the k th region of image j . Also, let $W_S = [w_k^s]_{k \in 1 \dots K}$ represent the relevance weights to be automatically learned.

To estimate the similarity between the ROI in the query image, and the database images, we compute the distance between the ROI i selected by the user and each region k in j as follows:

$$d_{ij} = \sum_{k=1}^K \left[\sum_{t=1}^S (W_{ik}^t)^2 (f_i^t - f_{jk}^t)^2 \right] \quad (1)$$

subject to

$$\sum_{t=1}^S W_{ik}^t = 1. \quad (2)$$

In order to learn W_{ik}^t , we minimize the distance d_{ij} in (1) by applying the Lagrangian function with linear constraints in (2) as follows:

$$L_{ij} = \sum_{k=1}^K \sum_{t=1}^S (W_{ik}^t)^2 (f_i^t - f_{jk}^t)^2 - \alpha_{ij} \left(\sum_{t=1}^S W_{ik}^t - 1 \right), \quad (3)$$

where α_{ij} is the lagrangian multiplier.

Next, in order to find the optimal weights, W_s , we set the derivative of L_{ij} with respect to W to zero:

$$\frac{\partial L}{\partial W_{ik}^t} = 2W_{ik}^t (f_i^t - f_{jk}^t)^2 - \alpha_{ij} = 0. \quad (4)$$

Thus,

$$W_{ik}^t = \frac{\alpha_{ij}}{2(f_i^t - f_{jk}^t)^2}. \quad (5)$$

By substituting W_{ik}^t in (2), we get:

$$\sum_{t=1}^S \frac{\alpha_{ij}}{2(f_i^t - f_{jk}^t)^2} = 1. \quad (6)$$

Thus,

$$\alpha_{ij} = \frac{1}{\sum_{t=1}^S \frac{1}{2(f_i^t - f_{jk}^t)^2}}. \quad (7)$$

Given (7) and (5), the optimal relevance weight W_{ik}^t are:

$$W_{ik}^t = \frac{D_{ik}^t}{\sum_{t=1}^S D_{ik}^t}, \quad (8)$$

where

$$D_{ik}^t = \left(\frac{1}{(f_i^t - f_{jk}^t)^2} \right)^2. \quad (9)$$

As it can be seen, the nominator of (8) represents the inverse of the distance between the features of the regions of i and k respectively. If this distance is large, the weight of the feature is low. On the other hand, if the distance is low, it means that the two regions are similar, so the corresponding weight is large. The overall proposed RBIR system that is based on these relevance feature weights is summarized in Algorithm 1.

Algorithm 1: (RBIR using relevance feature weights)

Input:

F_i^t : The t th feature of the query region i
 $\{F_{jk}^t\}_{j \in \{1, \dots, N\}}$: The t th feature of the k th region of the image j
 M : Number of retrieved image

Output: M retrieved images

1. For each image j ,
 - a. Compute the weight w using Eq. (8).
 - b. Compute the distance d_{ij} using Eq. (1).
 2. Sort $\{d_{ij}\}_{j \in \{1, \dots, N\}}$.
 3. Select the M images corresponding to the M lowest distances.
-

Table 1. Dataset statistics

Number of images	473
Number of extracted regions	2,631
Maximum number of region per image	19
Minimum number of region per image	1
Average number of regions per image	5
Number of labels	144

5. Experiments

In this section, we assess the performance of the proposed approach. In the following a description of the considered dataset, an outline of the retrieval performance measure used, and details about the performed experiments and results are provided.

5.1 Dataset

The dataset includes real images captured in various locations around the world. It is a collection of 475 pre-segmented images from different categories. More specifically, it is a subset of the segmented and annotated IAPR-TC12 benchmark (SAIAPR TC-12) [29]. SAIAPR TC-12 is an extension of the IAPR TC-12 benchmark [30] which was collected to assess the performance of automatic image annotation systems. IAPR TC-12 consists of 20,000 images collected from various locations around the world and includes categories such as “sports”, “cities”, “actions”, “animals”, “people”, etc. [31]. This subset of SAIAPR TC-12 benchmark represents the ground truth that we use to evaluate the performance of our proposed image retrieval approach. Figure 2 shows sample images from SAIAPR TC-12 benchmark. As it can be seen, the images belong to different categories.

These images are pre segmented into regions. The extracted regions are labeled. These labels represent the ground truth used to assess the retrieval performance. Table 1 reports the considered dataset statistics. Namely, it displays the number of images, number of considered regions, the maximum and minimum number of regions per image, average number of regions per image and number of labels.

5.2 Performance Measure

In order to assess the retrieval performance of our system, we use average normalized modified retrieval rank (ANMRR) [51]. It takes into consideration the order of the retrieved images.

Figure 2. Sample images from SAIAPR TC-12 benchmark [30].

ANMRR values are within the range[0, 1]. Low ANMRR value means highly accurate retrieval.

5.3 Experiment Description

We implement typical query based image retrieval as outlined in Section 2.1. First, as offline phase, the low-level features are extracted from the images to store their content in a database. These features are typically MPEG-7 low-level features [32]. Table 2 summarizes these features, their respective dimensionality and their corresponding parameters. Then, we apply two scenarios. The first one consists in using CBIR system for each extracted feature. The Euclidian distance is used to estimate the similarity between the query image and images from the database in the considered feature space. These distances are then sorted and the top M images corresponding to the lowest distances are displayed. The second scenario combines all the extracted features to represent the image content.

Besides, we assess the performance of the proposed RBIR. As offline phase, the features are extracted from all regions in order to represent them in the database. The size of the overall feature vector is 182. The online phase starts by extracting visual features from a given query region, then a relevance weight is assigned to each feature using Eq. (8) as described in Section 4. Next, the distance between the query region and all regions in the database are obtained using Eq. (1) as proposed in Section 4. Finally, these distances are sorted and the top M images corresponding to the lowest distances are retrieved.

Table 2. Features specification

Features	Dimensionality	Parameters	Color space
Color histogram visual feature	48	Number of levels= 16	RGB space
Color moment visual feature	9	-	HSV space
Color structure visual feature	64	Number of levels= 64	RGB space
Scalable color visual feature	32	Number of bins $H = 8$, Number of bins $S = 2$ and Number of bins $V = 2$	HSV space
Edge histogram visual feature	5	Threshold = 4	-
Wavelet transform visual feature	24	Number of Levels= 4	-

Table 3. Average retrieval scores ANMRR

Typical retrieval	0.47
Region-based retrieval	0.41

Figure 3. Average ANMMR score per label.

5.4 Experiment Results

Table 3 reports the retrieval results based on the ANMRR. As mentioned in Section 5.2, the lower the score is, the better the retrieval result is. As shown in Table 3, we notice that the proposed approach performs better than the typical retrieval system. In fact, the score is enhanced by 12.8%.

Figure 3 displays the average ANMMR score per label. We notice that although for few labels the typical retrieval is doing better than the proposed region-based approach (the blue line plot under the red line plot), for the majority of labels the region-based approach is doing better (the red line plot under the blue one).

Figure 4 displays the average ANMMR score per image. We notice that for the majority of the images, the proposed region-based approach is performing better than the typical approach since the average ANMMR score is lower for the region-based approach than for a typical retrieval approach.

For further investigation of the results, we display a sample retrieval result. Table 4, displays the retrieval results of the sample query image. This table shows the top 4 retrieved images of sample query images when using the proposed region-based retrieval approach, the typical retrieval approach using the different visual features separately, and feature the typical retrieval approach using an aggregation of all features. We should mention here that when using the typical retrieval approach the whole image is conveyed as input to the retrieval system. On the other hand, when using the proposed region-based approach,

Figure 4. Average ANMMR score per image.

only the part of the image containing the ROI is conveyed to the retrieval system. We notice from Table 4 that, when using the typical retrieval, the top 4 retrieved images are not relevant to sample query image 1 for any considered visual feature. On the other hand, the proposed approach is able to retrieve it correctly. In fact, we notice that the first retrieved image contains red flower, the second and the third retrieved images contain purple flowers, and the fourth one contains a red flower that has the same shape as the query region. Table 5 displays the relevance feature weights learned by the region-based retrieval system for the considered sample query image. We notice that for the first retrieved image, the relevant features are the SCD, CSD, and EHD. The two first ones combine both color and texture while the third one describes the edges of the region. Comparing the feature weights of the first and second retrieved images (Table 5), we notice that EHD is no more considered as relevant feature. On the other hand, the color moment becomes relevant. This is reflected in the retrieval results. In fact, we can see that the second retrieved image includes a purple flower of the same color as the query one but with a different texture.

As mentioned above, when using each feature separately the retrieval result is not relevant. However, when combining them with the appropriate relevant weights, the system is able to retrieve correctly the considered sample query image. Thus, by learning the appropriate relevance feature weights the proposed region-based approach captured the semantic meant by the user. This way, it is able to outperform the typical retrieval system.

6. Conclusion and Future Work

The main limitation of a typical query by example CBIR system is that the retrieval process takes into consideration the whole content of the query image. In particular, if it contains some irrelevant objects or regions, the retrieved images would include

Table 4. Top 4 retrieved images of the sample query images

	Query image	Retrieved image 1	Retrieved image 2	Retrieved image 3	Retrieved image 4
Region-based retrieval
Image-based retrieval using CH
Image-based retrieval using CM feature
Image-based retrieval using SCD feature
Image-based retrieval using CSD feature
Image-based retrieval using EHD feature
Image-based retrieval using DWT feature
Image-based retrieval using all

similar non-pertinent information. These irrelevant areas limit the accuracy of such CBIR systems. Moreover, these systems show other limitations due to the extraction of the visual features from the whole image, which yields a semantic gap between the image's high-level meaning and the extracted low-level features.

We presented a new region-based CBIR approach that associates relevance weights to the different low-level features in an unsupervised manner in order to better represent the visual

content of the image and match the user interest. We assessed the performance of the proposed approach by comparing its retrieval results to a typical retrieval system. The experimental results show that the proposed approach improved the retrieval results by 12.8%. Moreover, we further investigated the results and displayed a sample query image retrieval results and the corresponding relevance feature weights. We showed how the proposed approach captures the semantics of the user and translates it to feature weights. This yields the proposed approach to

Table 5. Relevance features weights learned by the region-based retrieval system for the sample query image

perform better than the typical retrieval approach.

As future work, we intend to consider rectangular block based regions of interest instead of irregular segmented one. We also plan to consider multiple regions of interest. We will then compare their corresponding retrieval results.

Conflict of Interest

No potential conflict of Interest relevant to this article was reported.

Acknowledgements

The authors are grateful for the support by the Research Center of the College of Computer and Information Sciences, King Saud University.

References

- [1] Apple, Available: <http://www.apple.com/>
- [2] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, B. Dom, et al., "Query by image and video content: the

- QBIC system,” *Computer*, vol. 28, no. 9, pp. 23-32, 1995. <https://doi.org/10.1109/2.410146>
- [3] J. R. Smith and S. F. Chang, “VisualSEEk: a fully automated content-based image query system,” in *Proceedings of the 4th ACM International Conference on Multimedia*, Boston, MA, 1997, pp. 87-98. <https://doi.org/10.1145/244130.244151>
- [4] J. Z. Wang, J. Li, and G. Wiederhold, “SIMPLIcity: Semantics-sensitive integrated matching for picture libraries,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 9, pp. 947-963, 2001. <https://doi.org/10.1109/34.955109>
- [5] E. R. Vimina and K. P. Jacob, “A sub-block based image retrieval using modified integrated region matching,” 2013, Available: <https://arxiv.org/abs/1307.1561>
- [6] F. Long, H. Zhang, and D. D. Feng, “Fundamentals of content-based image retrieval,” in *Multimedia Information Retrieval and Management*. Heidelberg: Springer, 2003, pp. 1-26. https://doi.org/10.1007/978-3-662-05300-3_1
- [7] A. A. Goodrum, “Image information retrieval: an overview of current research,” *Informing Science*, vol. 3, no. 2, pp. 63-66, 2000. <https://doi.org/10.28945/578>
- [8] N. K. Alham, M. Li, S. Hammoud, and H. Qi, “Evaluating machine learning techniques for automatic image annotations,” in *Proceedings of the 6th International Conference on Fuzzy Systems and Knowledge Discovery*, Tianjin, China, 2009, pp. 245-249. <https://doi.org/10.1109/FSKD.2009.531>
- [9] Y. Chen, J. Z. Wang, and R. Krovetz, “CLUE: cluster-based retrieval of images by unsupervised learning,” *IEEE Transactions on Image Processing*, vol. 14, no. 8, pp. 1187-1201, 2005. <https://doi.org/10.1109/TIP.2005.849770>
- [10] N. Varish and A. K. Pal, “Content based image retrieval using statistical features of color histogram,” in *Proceedings of the 3rd International Conference on Signal Processing, Communication and Networking*, Chennai, India, 2015, pp. 1-6. <https://doi.org/10.1109/ICSCN.2015.7219922>
- [11] J. Rose and M. Shah, “Content-based image retrieval using gradient projections,” in *Proceedings of the IEEE Southeastcon*, Orlando, FL, 1998, pp. 118-121. <https://doi.org/10.1109/SECON.1998.673306>
- [12] A. H. Kam, T. T. Ng, N. G. Kingsbury, and W. J. Fitzgerald, “Content based image retrieval through object extraction and querying,” in *Proceedings of IEEE Workshop on Content-based Access of Image and Video Libraries*, Hilton Head Island, SC, 2000, pp. 91-95. <https://doi.org/10.1109/IVL.2000.853846>
- [13] Y. M. Zhou, J. K. Wang, and A. M. Yang, “A method of region-based calculating image similarity for RBIR system,” in *Proceedings of the 9th International Conference for in Young Computer Scientists*, Hunan, China, 2008, pp. 814-819. <https://doi.org/10.1109/ICYCS.2008.248>
- [14] Y. Liu, D. Zhang, and G. Lu, “Region-based image retrieval with high-level semantics using decision tree learning,” *Pattern Recognition*, vol. 41, no. 8, pp. 2554-2570, 2008. <https://doi.org/10.1016/j.patcog.2007.12.003>
- [15] L. Yang, B. Geng, Y. Cai, A. Hanjalic, and X. S. Hua, “Object retrieval using visual query context,” *IEEE Transactions on Multimedia*, vol. 13, no. 6, pp. 1295-1307, 2011. <https://doi.org/10.1109/TMM.2011.2162399>
- [16] N. Amoda and R. K. Kulkarni, “Efficient image retrieval using region based image retrieval,” *Signal & Image Processing*, vol. 4, no. 3, pp. 17-29, 2013. <https://doi.org/10.5121/sipij.2013.4302>
- [17] F. Li and R. Liu, “Multi-graph multi-instance learning with soft label consistency for object-based image retrieval,” in *Proceedings of IEEE International Conference on in Multimedia and Expo*, Turin, Italy, 2015, pp. 1-6. <https://doi.org/10.1109/ICME.2015.7177391>
- [18] K. Vu, K. A. Hua, and W. Tavanapong, “Image retrieval based on regions of interest,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 15, no. 4, pp. 1045-1049, 2003. <https://doi.org/10.1109/TKDE.2003.1209021>
- [19] S. Wei, Y. Zhao, and Z. Zhu, “Multilevel indexing structure for object based image retrieval,” in *Proceedings of the 8th International Conference in Signal Processing*, Beijing, China, 2006. <https://doi.org/10.1109/ICOSP.2006.345575>
- [20] Y. Wang, K. B. Jia, and P. Y. Liu, “A novel ROI based image retrieval algorithm,” in *Proceedings of the 2nd International Conference on Innovative Computing, Information and Control*, Kumamoto, Japan, 2007. <https://doi.org/10.1109/ICICIC.2007.69>

- [21] Y. Velazco-Paredes, R. Flores-Quipe, and R. E. Patino Escarcina, "Region-based image retrieval using color and texture features on irregular regions of interest," in *Proceedings of IEEE Colombian Conference on Communications and Computing*, Popayan, Colombia, 2015, pp. 1-6. <https://doi.org/10.1109/ColComCon.2015.7152077>
- [22] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, no. 2, pp. 91-110, 2004. <https://doi.org/10.1023/B:VISI.0000029664.99615.94>
- [23] K. J. Cios, W. Pedrycz, and R. W. Swiniarski, *Data Mining Methods for Knowledge Discovery*. Boston, MA: Springer, 1998.
- [24] T. Chang and C. C. J. Kuo, "Texture analysis and classification with tree-structured wavelet transform," *IEEE Transactions on Image Processing*, vol. 2, no. 4, pp. 429-441, 1993. <https://doi.org/10.1109/83.242353>
- [25] A. Laine and J. Fan, "Texture classification by wavelet packet signatures," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 15, no. 11, pp. 1186-1191, 1993. <https://doi.org/10.1109/34.244679>
- [26] N. Beckmann, H. P. Kriegel, R. Schneider, and B. Seeger, "The R*-tree: an efficient and robust access method for points and rectangles," *ACM SIGMOD Record*, vol. 19, No. 2, pp. 322-331, 1990. <https://doi.org/10.1145/93605.98741>
- [27] K. A. Hua, K. Vu, and J. H. Oh, "SamMatch: a flexible and efficient sampling-based image retrieval technique for large image databases," in *Proceedings of the 7th ACM International Conference on Multimedia (Part 1)*, Orlando, FL, 1999, pp. 225-234. <https://doi.org/10.1145/319463.319610>
- [28] N. Jhanwar, S. Chaudhuri, G. Seetharaman, and B. Zavidovique, "Content based image retrieval using motif cooccurrence matrix," *Image and Vision Computing*, vol. 22, no. 14, pp. 1211-1220, 2004. <https://doi.org/10.1016/j.imavis.2004.03.026>
- [29] H. J. Escalante, C. A. Hernandez, J. A. Gonzalez, A. Lopez-Lopez, M. Montes, F. F. Morales, L. Enrique Sucar, L. Villaseñor, and M. Grubinger, "The segmented and annotated IAPR TC-12 benchmark," *Computer Vision and Image Understanding*, vol. 114, no. 4, pp. 419-428, 2010. <https://doi.org/10.1016/j.cviu.2009.03.008>
- [30] ImageCLEF, "IAPR TC-12 Benchmark," Available: <http://imageclef.org/photodata>
- [31] M. Grubinger, "Analysis and evaluation of visual information systems performance," PhD dissertation, Victoria University, Melbourne, Australia, 2007.
- [32] B. S. Manjunath, P. Salembier, and T. Sikora, *Introduction to MPEG-7: Multimedia Content Description Interface*. Chichester: John Wiley & Sons, 2002.

Ouiem Bchir is an associate professor at the Department of Computer Science, College of Computer and Information Sciences (CCIS), King Saud University, Riyadh, Saudi Arabia.

Dr. Ouiem Bchir obtained her Ph.D. from the University of Louisville, KY, USA. Her research interests are spectral and kernel clustering, pattern recognition, hyper-spectral image analysis, local distance measure learning, and unsupervised and semi-supervised machine learning techniques. She received the University of Louisville Dean's Citation, the University of Louisville CSE Doctoral Award, and the Tunisian presidential award for the electrical engineering diploma.

E-mail: ouiem.bchir@gmail.com

Mohamed Maher Ben Ismail is an associate professor at the Department of Computer Science of the College of Computer and Information Sciences at King Saud University. He received his Ph.D. degree in Computer Science from the University of Louisville in 2011. His research interests include Pattern Recognition, Machine Learning, Data Mining and Image Processing.

E-mail: maher.benismail@gmail.com

Hadeel Aljam got her master degree in computer science from King Saud University, Riyadh, Saudi Arabia.

E-mail: obchir@ksu.edu.sa

Photo is not included by the author's request.