

RESEARCH

Open Access



Taking advantage of motif matrix inference for rotated image indexing and retrieval

Yi Xu^{1,2}, Wei Song^{1,3*}, Yu Yan¹, Huadong Zhu¹ and Ting Cao¹

Abstract

With the rapid development of information technology, the sizes of digital libraries become larger and larger. How to quickly and effectively search the desired images in huge digital libraries becomes the challenge needed to resolve with high priority. In this study, we firstly propose two motif-based matrices, i.e., the motif average matrix (MAM) and motif excessive matrix (MEM), to describe the color and texture features of an image. Subsequently, in terms of the inference of MAM and MEM, a motif matrix (MM) is further proposed to index rotated images and resolve the issue of rotated image retrieval. That is, in the light of such an inference, MM incorporates the color and texture characters and reveals the consistent relevance between the original and rotated images, which can be effectively used for rotated image retrieval. To extensively test the performance of our method, we carry out the experiments on the benchmark Corel image dataset, Brodatz texture image dataset, WIPO global brand dataset, and the experimental results show that our approach of motif matrix inference improves the retrieval performance in comparison with the state-of-the-art image retrieval approaches.

Keywords: Content-based image retrieval, Image indexing and presentation, Motif matrix, Rotated images analysis

1 Introduction

In recent years, with the rapid development of multimedia and computer technologies, the sizes of digital libraries become larger and larger. As a result, in these large digital libraries, how to find desired image information, especially for numerous rotate images, has become a challenge needed to resolve urgently. Rotated image retrieval has a wide range of applications in IR systems. As the number of trademark images is increasing rapidly in trademark registration system, the design of new patterns should prevent the conflict to the similar trademarks registered. Especially, the similar patterns caused by rotation also need to be effectively avoided. Trademark image retrieval system can find similar trademarks immediately after entering a new trademark image for registration, which can effectively protect the legitimate rights and interests of registered trademarks. In recent decade, image retrieval has become a research hot spot

in the fields of image processing, pattern recognition, and artificial intelligence.

Image retrieval technologies based on text [1, 2] and content [3, 4] have been extensively studied in the last decade. Text-based image retrieval approach is firstly proposed and widely utilized in the 1970s. Large search engine companies, e.g., Google, Yahoo, and Baidu, extensively use keywords annotated on images to implement retrieval. However, with the widely spread digital imaging devices, such a kind of approach has two drawbacks. Firstly, the need of manually handling image databases can be too expensive. Secondly, the results of retrieval may be inaccurate, because they are frequently related to the subjective understanding of annotators. To distinguish from text-based approach, content-based image retrieval (CBIR) [5–8] has been proposed in the early 1990s. This approach is to retrieve images using low-level features like color [9–11], shape [12, 13], and texture [14–16], to describe an image. Through extracting the natural features of an image, CBIR calculates the similarity between the query image and database images [17, 18] and ranks the mostly related images in terms of the similarities. Such an approach completely releases

* Correspondence: songwei@jiangnan.edu.cn

¹School of Internet of Things (IoT) Engineering, Jiangnan University, Wuxi, China

³Jiangsu Provincial Engineering Laboratory of Pattern Recognition and Computational Intelligence, Jiangnan University, Wuxi, China

Full list of author information is available at the end of the article

the labor of annotators and effectively reduces inaccurate feedbacks without the subjective depiction [19–21]. Moreover, a large number of CBIR systems have been developed by various organizations, individuals, and hospitals, e.g., QBIC [22], PhotoBook [23], VisualSEEK [24], Netra [25], Pictoseek [26], SIMPLIcity [27], and Blobworld [28].

As an important topic in the research field of image retrieval, CBIR is extensively studied. Generally speaking, CBIR usually utilizes descriptors, i.e., color, texture, or shape, to represent an image. Various algorithms have been designed to extract such features. HSV histogram (HSVH) [29] extracts the color pixel features by HSV color space. It transfers each image to a quantized color histogram for subsequent image comparison. In this scheme, the color feature is just taken into account, and the spatial distribution feature of pixels is ignored. Color co-occurrence matrix (CCM) [30] is a conventional pattern co-occurrence matrix that calculates the probability of the occurrence of same arrangement between each pair of adjacent motifs, and this probability is considered as the attribute of the image. On the basis of CCM, a modified color motif co-occurrence matrix (MCMCM) [31] is given to collect the inter-correlation between the red, green, and blue color space which is absent in CCM. However, the intensity of pixel variation in each motif is not considered in these two methods and usually causes the value of the 2×2 pixels which have large differences share the same motif. In view of such a defect, difference between pixels of scan pattern (DBPSP) [32] is presented and intended to calculate the differences among all pixels within motifs. The combined algorithm, named MCMCM&DBPSP [33], respectively calculates the similarity obtained from MCMCM and DBPSP first and then utilizes a fixed coefficient to normalize the weights of them. However, it is overly dependent on the predefined coefficient which is difficult to select, especially for applications to different image datasets. Structure elements' descriptor (SED) [34] is a kind of a texture descriptor based on HSV color space. It extracts color and texture features to represent an image. However, it is very sensitive to the image whose regions or textures are not significant, and the effectiveness is restricted. Moreover, the issue of image rotation is not considered in the above methods which would generate error of judgment. Thus, the retrieval results seem to contradict human intuition for recognizing similar images. In the research fields of fingerprint identification and LOGO registration, many rotation invariant features or models, e.g., Gabor transformation and statistical features [35], wavelet hidden Markov model [36], and local binary pattern [37], are taken into account. However, the large amount of computations and the

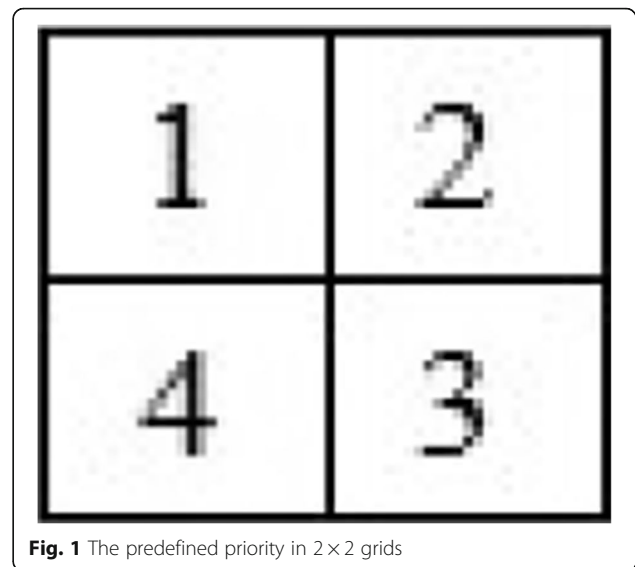


Fig. 1 The predefined priority in 2×2 grids

high dimensionality of features usually affect the performance of retrieval.

In this paper, we proposed a motif matrix inference (MMI) based on rotation invariant texture features for rotated image retrieval. To tackle with the issue of rotated image retrieval, we implement a minimal upper left triangle (MULT) rule which concludes the final eight motifs to keep the consistence on the basis of analysis of statistics. Such an approach can effectively reduce the kind of motifs and resolve the issue of distortion after image rotation. A motif excessive matrix (MEM) is subsequently derived from the motif transformed image to depict a whole image. The whole image is divided into 2×2 pixel grids and each grid is replaced by a scan motif. Meanwhile, we utilize motif average matrix (MAM) to store the gray information of an image. Finally, a motif matrix (MM) is further proposed to integrate MEM and MAM to extract and describe color and texture features. That is, we can obtain the same MM when an image is rotated. Therefore, our approach integrates the advantages of both textural and colorful description methods. What is more, through the analysis of statistics, the proposed approach using MULT reduces the space of MM and effectively solves the issue of rotated image retrieval.

The remainder of the paper is structured as follows. Section 2 gives the definition of motif and introduces how to transform an image into its relevant MAM, MEM, and MM. The original MM with the size of 256×24 only counts the textural and gray information of an image. Its inherent structure cannot resolve the issue of rotated image retrieval. Meanwhile, the simply merging motif may cause the problem of

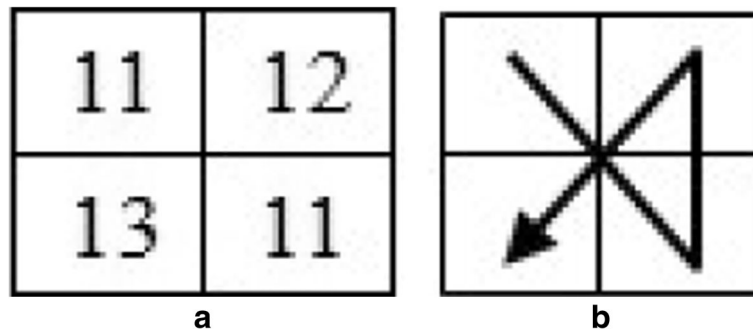


Fig. 2 **a** The original gray of each grid. **b** The unique motif in 2×2 grids

image distortion. In this study, the proposed MULT reduces the space of MM into a relatively smaller scale of 256×8 and can effectively resolve the issue of rotated image retrieval. Section 3 describes the evaluation method to calculate the similarity between images. The experimental results and comparisons are presented in Section 4. Conclusion is given in Section 5.

2 Methods

2.1 The original motif

In this paper, the original image is divided into 2×2 grids. These grids are then replaced by a particular scan motif which would traverse the grid in the ascending gray order (AGO), to reflect the texture of 2×2 grids (see Fig. 1). Note that a fixed priority is given to resolve

NO.	0	1	2	3	4	5
motif						
sequence	1234	1243	1324	1342	1423	1432
NO.	6	7	8	9	10	11
motif						
sequence	2134	2143	2314	2341	2413	2431
NO.	12	13	14	15	16	17
motif						
sequence	3124	3142	3214	3241	3412	3421
NO.	18	19	20	21	22	23
motif						
sequence	4123	4132	4213	4231	4312	4321

Fig. 3 Total 24 kinds of motifs

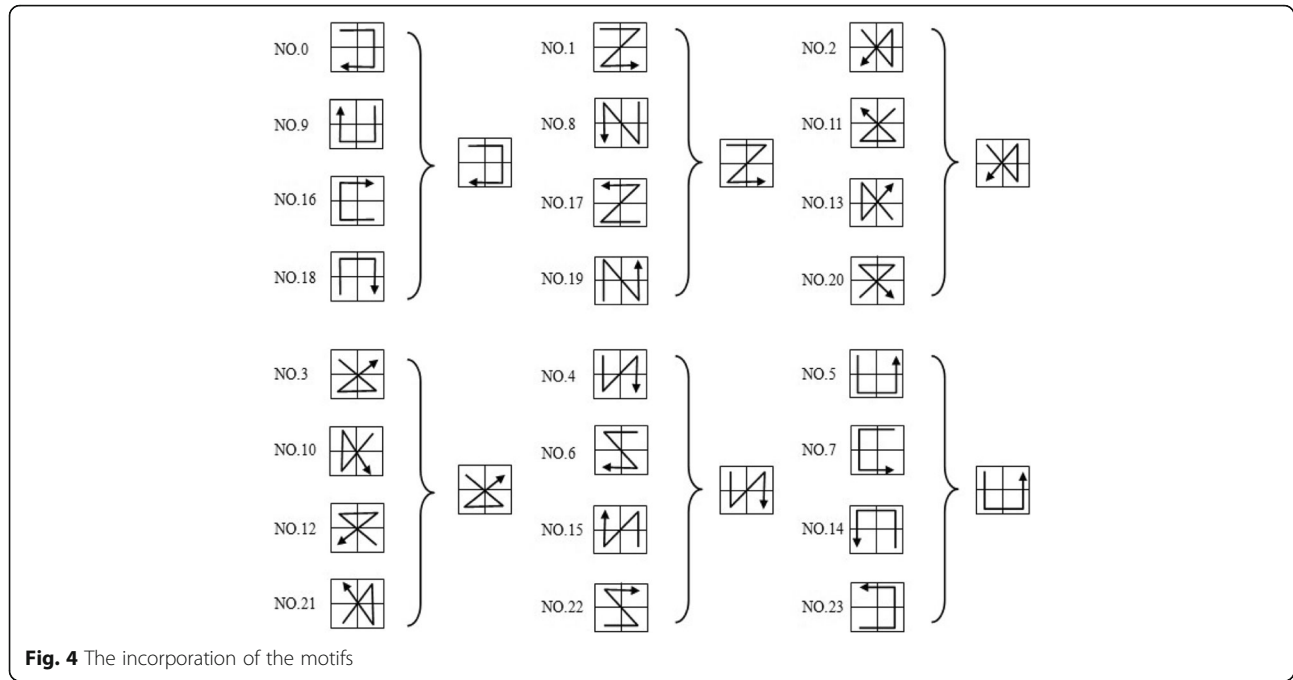


Fig. 4 The incorporation of the motifs

the issue of the grids that have same gray. That is, in terms of such a priority, a unique scan curve can be achieved to represent the 2×2 grids.

In Fig. 1, we define the upper left grid (label 1) has the highest priority and the bottom left grid (label 4) has the lowest priority. Such priority order can solve the problem of linking curve with the same gray for grids. Meanwhile, with the ascending gray order, we can transform 2×2 grids into a unique motif. Take Fig. 2 for instance, although the gray of label 1 is same to that of label 3, we can also obtain a unique motif with sequence 1324. In general, 24 different motifs could traverse “ 2×2 grids” (see Fig. 3).

2.2 The modified eight rotated invariant motifs

It is intuitive that if a motif rotates 90° , 180° , and 270° clockwise it will transform to another three kinds of motifs respectively. Therefore, to handle rotated image retrieval, we classified the original motif and its three rotated motifs as one motif. That is, after such direct incorporation, we can see from Fig. 4 that the total 24

kinds of motifs can be directly reduced to 6 motifs, and the new 6 kinds of motifs are shown in Fig. 5.

It seems that such a simple incorporation process may solve the rotated image retrieval. However, after further analysis, it may cause the issue of rotated distortion. That is, if a unit of 2×2 grids encounters the situation of the same grays, the rotation of the original 2×2 grids may cause the different motifs out of the incorporated motif assigned. We take Fig. 6 for instance. Through the simple incorporation, the motif related to the original grids is assigned to the kind of NO.2 motif, but the rotated grids via the three degrees produce the NO.0 kind motif, which is different from that produced by the original grids and causes the issue of rotated distortion. Thus we cannot just utilize such a simple incorporation to deal with rotated image retrieval. In this study, we propose a rule of minimal upper left triangle (MULT) to solve it, and the MULT is described as:

- Step 1: Calculate the sum of the grays in labels 1, 2, and 4.
- Step 2: If the sum is less than or equal to that of other three adjacent grays, i.e., labels 1, 2, and 3; labels 1, 3, and 4; labels 2, 3, and 4, go to Step 4. Otherwise, go to Step 3.
- Step 3: Rotate the 2×2 grids via 90° clockwise and calculate the sum of the grays in labels 1, 2, and 4 again, then go to Step 2.
- Step 4: Utilize AGO to obtain the motif.

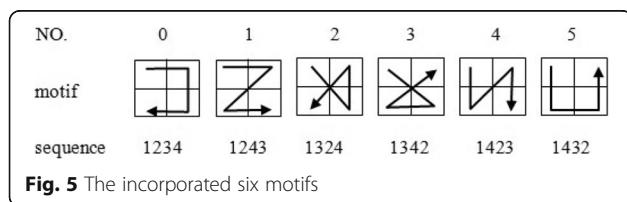


Fig. 5 The incorporated six motifs

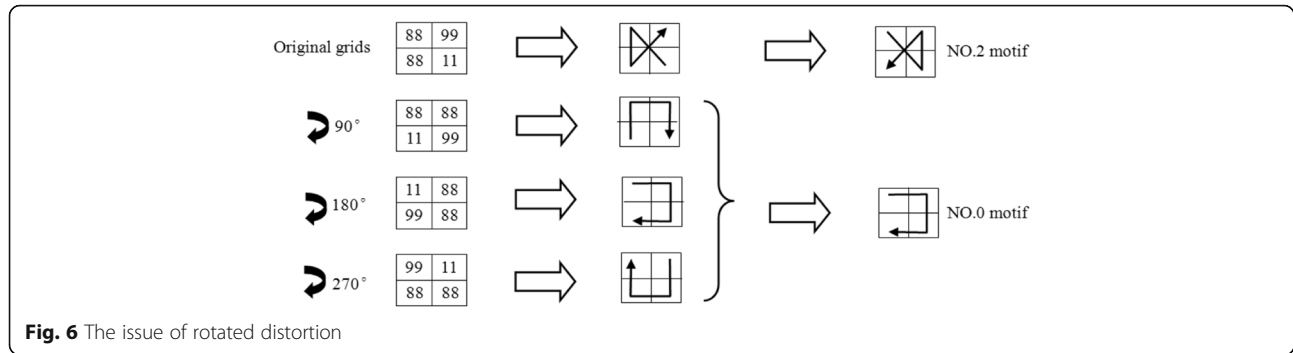


Fig. 6 The issue of rotated distortion

The MULT rule fixed the three smaller grays in labels 1, 2, and 4 grids, no matter which degree (90°, 180°, or 270°) is taken for rotating the original image. In Fig. 7, we also take the 2×2 grids (see Fig. 10) for example. Although the unit of 2×2 grids encounters the situation of the same grays, a unique motif can be finally achieved by the MULT rule.

According to the MULT rule, the original 24 kinds of motifs will be finally reduced to 8 kinds of motifs by a complete statistics method, and these 8 kinds of motifs are shown in Fig. 8.

It has been known that for each unit of 2×2 grids, the rotation of an image not only moves its position in space, but also rotates its inner content. After the further analysis and extensive simulations for MULT, we find the two special cases which are given by.

Case 1: A unit of 2×2 grids with three equivalent grays and another smaller gray.

Case 2: A unit of 2×2 grids with two equivalent grays and another two smaller grays.

To well handle these two cases, we improve MULT and specify them with one of its generated motif. The specifying process is illustrated in Fig. 9.

In order to discuss why we select these three kinds of generated motifs to represent the finally incorporated motifs, we firstly investigate the frequency of each case and motif in commonly used image dataset Core-1000 [38–40], and the corresponding frequency and percentage are shown in Tables 1 and 2.

Note that we need to select some distinguished motifs to represent the finally incorporated motif with respect to these two special cases. That is, it is intuitive that the generated motif with the lowest percentage would be selected. Subsequently, we will explain why we select these three kinds of generated motifs in Fig. 9 as follows.

- (1) For Case 1, from the original grid and the corresponding three rotated grids, three kinds of motifs, i.e., NO.0, NO.5, and NO.7, are generated with percentage of 11.787%, 10.31%, and 0% respectively. Thus, the motif of NO.7 is selected to represent Case 1.
- (2) For Case 2.1, from the original grid and the corresponding three rotated grids, two kinds of motifs, i.e., NO.4 and NO.6, are generated with percentage of 4.596% and 3.882% respectively. Thus, the motif of NO.6 is selected to represent Case 2.1.

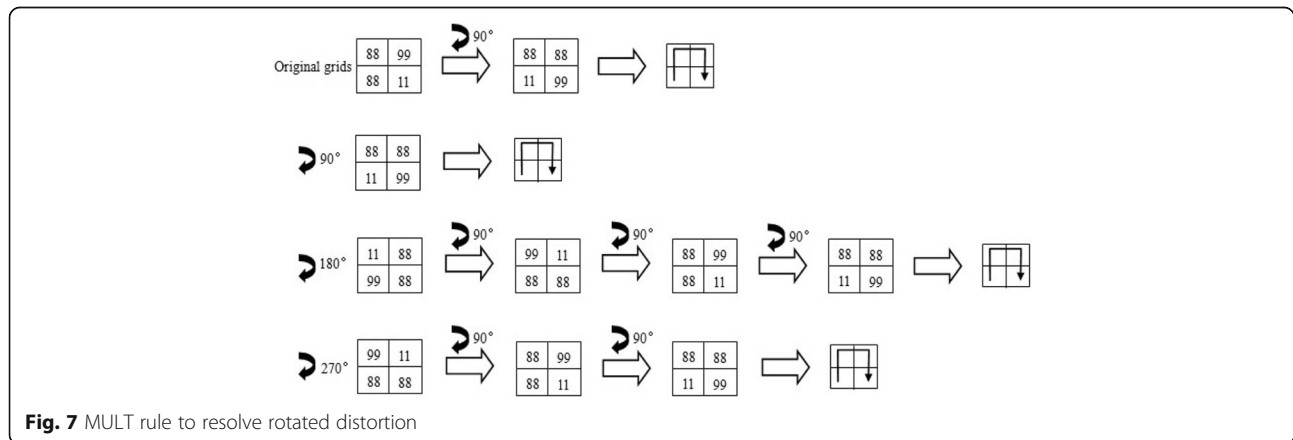


Fig. 7 MULT rule to resolve rotated distortion

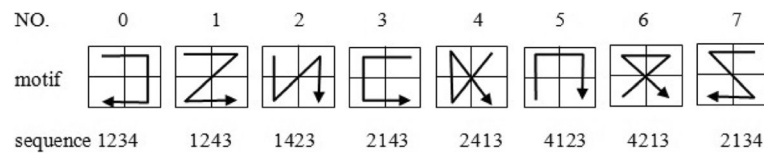
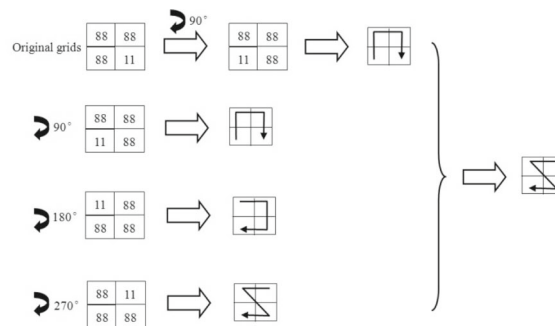


Fig. 8 The final eight motifs

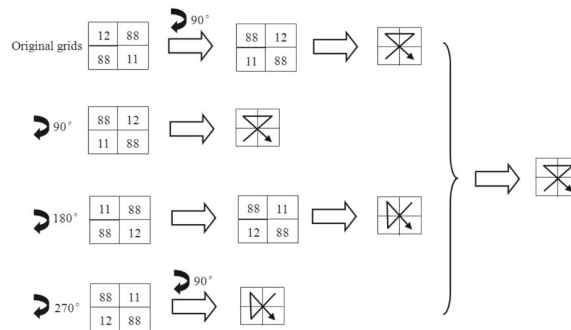
(3) For Case 2.2, from the original grid and the corresponding three rotated grids, two kinds of motifs, i.e., NO.2 and NO.7, are generated with percentage of 17.901% and 0 respectively. Thus, the motif of NO.7 is selected to represent Case 2.2.

So far, after the above modification for MULT, the number of motifs has been incorporated into the final eight motifs and such an incorporated approach can well organize MEM and MAM for rotated image retrieval.

Case 1:



Case 2.1:



Case 2.2:

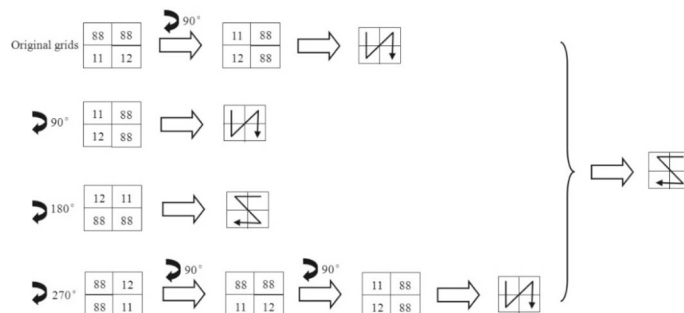


Fig. 9 The direct incorporation of the two special cases with its generated motif

Table 1 The frequency and percentage of each case in Corel-1000 image dataset

	Case		
	1	2.1	2.2
Frequency	2,575,059	858,361	12,614,285
Percentage	2.637	0.879	12.916

2.3 The motif average matrix (MAM)

In order to extract the color feature of an image, we utilize MAM to store the average gray in each motif. That is, since every element is obtained from 2×2 grids, it is obviously known that the gray level matrix with respect to a size of $M \times N$ image can be transformed into $(M - 1) \times (N - 1)$ MAM, where M and N are the pixel numbers of rows and columns in original image respectively. The steps to construct MAM are described as follows:

- (1) Starting from the origin location (0, 0) in gray level matrix and moving every unit of 2×2 grids from left to right and from top to bottom with 1-step length.
- (2) Extracting the average gray in each unit of 2×2 grids as the element of MAM.

We can take Figs. 10 and 11 for instance, Fig. 10 denotes an original 10×10 Gy level matrix for image and Fig. 11 denotes the related 9×9 MAM.

2.4 The motif excessive matrix (MEM)

In this study, both characters of color and texture are considered to comprehensively describe an image. That is to say, we not only use MAM to store the color features, but also utilize MEM to preserve the texture features. Since each element in MEM is a motif, and we extract the motif from each unit of 2×2 grids, it is obvious that the size of MEM is also $(M - 1) \times (N - 1)$ matrix associated with a $M \times N$ gray level image. Note that, the motif and the average gray information are all extracted from each unit of 2×2 grids. Thus, the texture feature in MEM can be mapped to the color feature in MAM with the related location. The steps to construct MEM are described as follows:

- (1) Starting from the origin location (0, 0) in gray level matrix and moving every unit of 2×2 grids from left to right and from top to bottom with one-step length.
- (2) Utilizing the method of AGO to generate the motif as the element of MEM.

Take Figs. 4 and 6 for instance. Based on the original gray level matrix (see Fig. 4), we can gain the relevant 9×9 MEM in Fig. 12.

2.5 The motif matrix (MM)

As discussed above, MAM just calculates the color feature, while MEM just counts the texture feature. In this study, the texture and color features are both considered as the content of an image. That is, we propose MM to further fuse these two features into a space. It has been known that the value of gray ranges from 0 to 255, and we have obtained total eight kinds of motifs to represent the textures. Subsequently, we utilize the range of gray and the kind of motifs to represent the rows and columns of MM respectively. That is to say, in the new generated MM is a 256×8 matrix, the element $MM(x, y)$ counts the times where in the same pixel (i, j) , $MAM(i, j)$ is the average gray of x , and $MEM(i, j)$ is the y th motif. The definition of element $MM(x, y)$ is given by

$$\begin{aligned} &\text{If } MAM(i, j) = x, x \in [0, 255] \\ &\text{and } MEM(i, j) = y, y \in [0, 7] \\ &\text{Then } MM(x, y) = MM(x, y) + 1 \end{aligned} \quad (1)$$

For example, if $MAM(20, 5)$ is 255 and $MEM(20, 5)$ is the second motif, the value of $MM(255, 2)$ should increase to 1. Through such an approach, we merge the texture and color features in a fused space. That is, MM counts the total times where each kind of motif shares the same average gray. Therefore, such a statistics-based approach can resolve the issue of searching the similar texture and color features which may disperse in a different position of an image.

Take the MAM in Fig. 11 and MEM in Fig. 12 for example. We can obtain the fused MM with a size of 256×8 . In order to explain MM in a clear way, we just take the 71st row of MM into account. That is, we count the times of each kind of motif whose average gray is 70. We can see from Fig. 11 that $MAM(0,6) = 70$, $MAM(1,1) = 70$,

Table 2 The frequency and percentage of each motif in Corel-1000 image dataset

	#Motif							
	0	1	2	3	4	5	6	7
Frequency	11,511,792	24,205,880	17,483,214	10,066,814	4,489,113	10,068,883	3,791,599	0
Percentage	11.787	24.785	17.901	10.307	4.596	10.310	3.882	0

25	7	29	76	95	29	45	12	62	251
211	129	55	96	81	109	192	30	210	12
41	2	92	100	208	87	199	207	0	17
88	37	129	217	92	207	12	103	19	8
11	2	84	42	66	95	59	33	37	71
7	9	49	77	5	40	240	255	77	37
102	88	52	100	100	80	40	27	32	2
88	72	25	200	211	30	72	54	42	49
207	21	107	0	107	82	96	106	79	9
255	207	89	209	87	33	69	7	219	100

Fig. 10 An original 10 × 10 Gy level matrix

$MAM(5,2) = 70$, $MAM(7,5) = 70$, $MAM(7,7) = 70$, $MAM(8,5) = 70$, $MAM(8,6) = 70$, and from Fig. 12 $MEM(0,6) = 1$, $MEM(1,1) = 4$, $MEM(5,2) = 2$, $MEM(7,5) = 1$, $MEM(7,7) = 2$, $MEM(8,5) = 2$, $MEM(8,6) = 5$, we utilize Table 3 to depict the 71st row of MM.

In this study, we focus on CBIR, especial for rotated image retrieval. So far we have utilized MM to organize the color and texture features of an image. To clearly depict the issue of image rotation, we also take the original 10 × 10 Gy level matrix from Fig. 10 for instance.

After rotation of 90°clockwise, we can gain the rotated gray level matrix, and the relevant MAM and MEM shown in Fig. 13a–c respectively. Subsequently, the new generated MAM and MEM are utilized to construct the relevant MM. In order to compare this new constructed MM with the original MM without rotation, we also take the 71st row of new MM into account. We can see from Fig. 13b that $MAM(1,7) = 70$, $MAM(2,3) = 70$, $MAM(5,0) = 70$, $MAM(5,1) = 70$, $MAM(6,0) = 70$, $MAM(6,8) = 70$, and $MAM(7,1) = 70$ and from Fig. 13c $MEM(1,7) = 4$, $MEM(2,3) = 2$, $MEM(5,0) = 2$, $MEM(5,1) = 1$, $MEM(6,0) = 5$, $MEM(6,8) = 1$, and $MEM(7,1) = 2$. Table 4 is used to depict the 71st row of the new MM.

93	55	64	87	79	94	70	79	134
96	70	86	121	121	147	157	112	60
42	65	135	154	149	126	130	82	11
35	63	118	104	115	93	52	48	34
7	36	63	48	52	109	147	101	56
52	50	70	71	56	100	141	98	37
88	59	94	153	105	56	48	39	31
97	56	83	130	108	70	82	70	45
173	106	101	101	77	70	70	103	102

Fig. 11 The constructed MAM

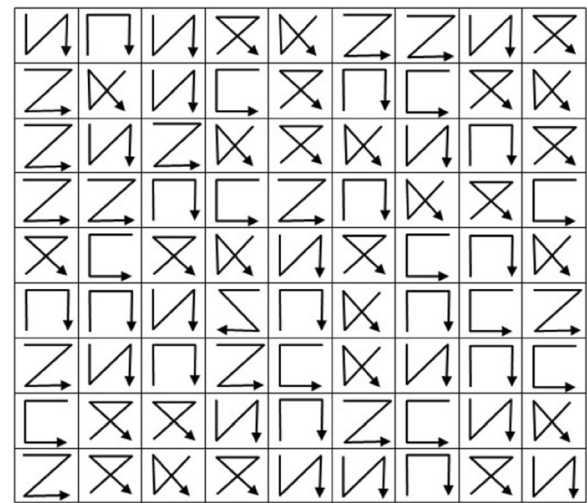


Fig. 12 The constructed MEM

In comparison with Table 3, we can see that the 71st row of new MM in Table 4 is obviously the same with that of the original MM. That is, through using MM and the eight kinds of modified motifs, we can resolve the issue of image rotation.

3 Similarity measure

In this study, the MM is proposed to depict the texture and color features of an image. Thus, we can only calculate the distance between MMs to reveal the similarity between relevant images. For each template image in the database, its MM with the size of 256×8 has been extracted and stored in advance. In this paper, the distance between the elements related to the template MM and query MM is given by

$$D(MM_T(i, j), MM_Q(i, j)) = \frac{|MM_T(i, j) - MM_Q(i, j)|}{1 + MM_T(i, j) + MM_Q(i, j)} \quad (2)$$

where $MM_T(i, j)$ and $MM_Q(i, j)$ are the respective element of template MM and query MM. Since $MM(i, j)$ counts the times of the i th kind of motif whose average gray is j , it is actually set as a positive integer or zero. It is intuitive that when $|MM_T(i, j) - MM_Q(i, j)| \rightarrow 0$, $D(MM_T(i, j), MM_Q(i, j)) \rightarrow 0$. That is, the similar features would lead to the shorter distance. On the contrary, as $|MM_T(i, j) - MM_Q(i, j)| \rightarrow \infty$, $D(MM_T(i, j), MM_Q(i, j)) \rightarrow 1$. That is, the greater contrast would

Table 3 The 71st row in MM obtained from Figs. 11 and 12

#Motif	0	1	2	3	4	5	6	7
Times	0	2	3	0	1	1	0	0

a

255	207	88	102	7	11	88	41	211	25
207	21	72	88	9	2	37	2	129	7
89	107	25	52	49	84	129	92	55	29
209	0	200	100	77	42	217	100	96	76
87	107	211	100	5	66	92	208	81	95
33	82	30	80	40	95	207	87	109	29
69	96	72	40	240	59	12	199	192	45
7	106	54	27	255	33	103	207	30	12
219	79	42	32	77	37	19	0	210	62
100	9	49	2	37	71	8	17	12	251

The 10 * 10 gray level matrix rotated by 90°clockwise from Fig. 10.

b

173	97	88	52	7	35	42	96	93
106	56	59	50	36	63	65	70	55
101	83	94	70	63	118	135	86	64
101	130	153	71	48	104	154	121	87
77	108	105	56	52	115	149	121	79
70	70	56	100	109	93	126	147	94
70	82	48	141	147	52	130	157	70
103	70	39	98	101	48	82	112	79
102	45	31	37	56	34	11	60	134

The generated MAM.

c

↗	↘	↗	↘	↗	↘	↗	↘	↗	↘
↘	↗	↘	↗	↘	↗	↘	↗	↘	↗
↗	↘	↗	↘	↗	↘	↗	↘	↗	↘
↘	↗	↘	↗	↘	↗	↘	↗	↘	↗
↗	↘	↗	↘	↗	↘	↗	↘	↗	↘
↘	↗	↘	↗	↘	↗	↘	↗	↘	↗
↗	↘	↗	↘	↗	↘	↗	↘	↗	↘
↘	↗	↘	↗	↘	↗	↘	↗	↘	↗
↗	↘	↗	↘	↗	↘	↗	↘	↗	↘
↘	↗	↘	↗	↘	↗	↘	↗	↘	↗

The generated MEM.

Fig. 13 The rotated image and its generated MAM and MEM. **a** The 10 × 10 gray level matrix rotated by 90°clockwise from Fig. 10. **b** The generated MAM. **c** The generated MEM.

result in the bigger distance. Thus, the distance between the template MM and the query MM is given by

$$D_{MMI}(MM_T, MM_Q) = \frac{1}{\sum_{i=0}^{255} \sum_{j=0}^7 D(MM_T(i, j), MM_Q(i, j))} \quad (3)$$

where MM_T and MM_Q are respective motif matrix of template image and query image. Subsequently, the similarity between the template image and the query image is defined as

$$S_{MMI}(T, Q) = \frac{1}{1 + D_{MMI}(MM_T, MM_Q)} \quad (4)$$

The above formula transforms the similarity via conventional cognition. That is, the shorter distance makes

Table 4 The 71st row in MM obtained from Fig. 13b, c

#Motif	0	1	2	3	4	5	6	7
Times	0	2	3	0	1	1	0	0

the bigger similarity and vice versa. Moreover, it can avoid the problem of denominator being 0. In the field of image retrieval, Euclidean distance and Cosine similarity are usually adopted as the evaluation methods; the respective formula is given by

$$D_{Euc}(MM_T, MM_Q) = \sqrt{\sum_{i=0}^{255} \sum_{j=0}^7 (MM_T(i, j) - MM_Q(i, j))^2} \quad (5)$$

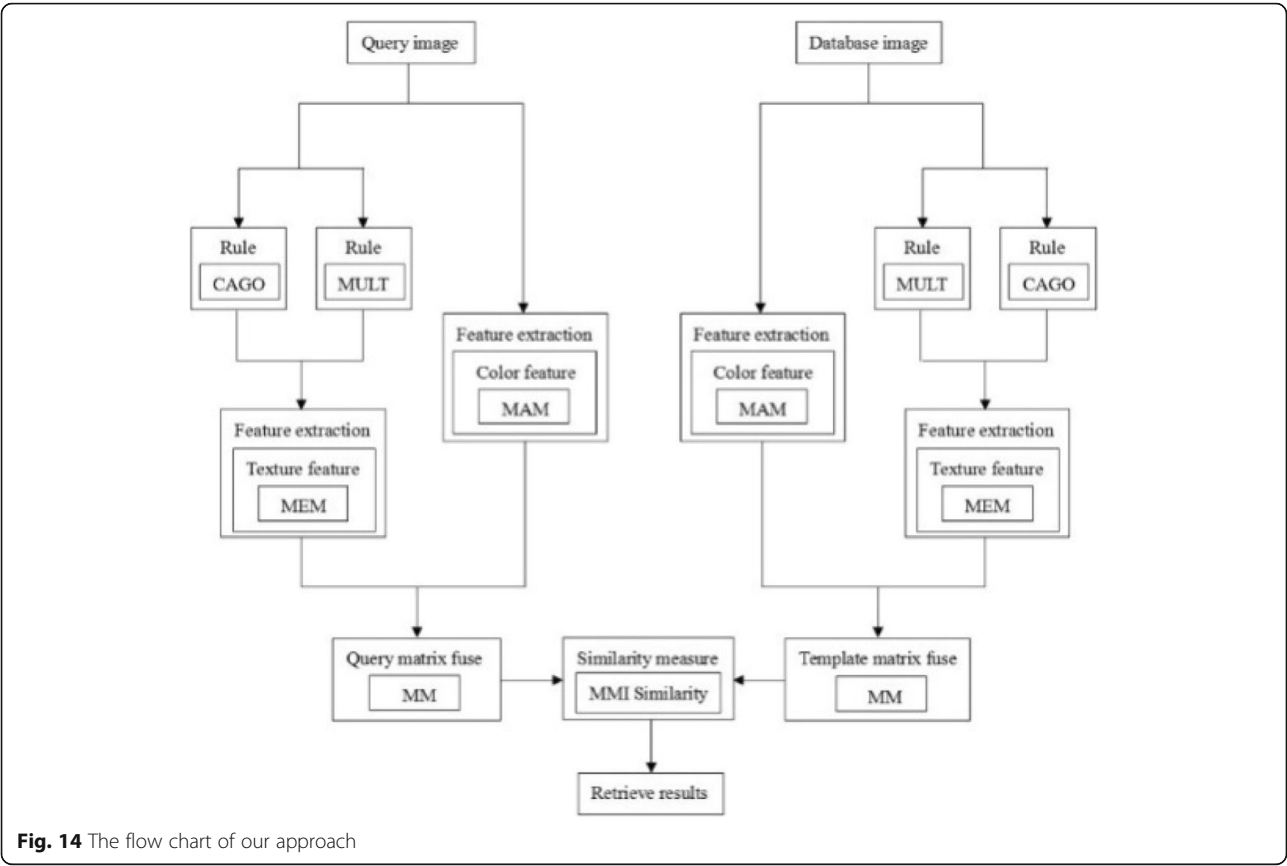
$$S_{Cos}(MM_T, MM_Q) = \frac{\sum_{i=0}^{255} \sum_{j=0}^7 MM_T(i, j) * MM_Q(i, j)}{\sqrt{\sum_{i=0}^{255} \sum_{j=0}^7 MM_T(i, j)^2} * \sqrt{\sum_{i=0}^{255} \sum_{j=0}^7 MM_Q(i, j)^2}} \quad (6)$$

In comparison with Euclidean distance and Cosine similarity, the evaluation measure given by us is easy to calculate the similarity without square or square root operation. Moreover, our approach is more effective than them, since it can achieve a better performance compared to those of the relatively more complicated Euclidean distance and Cosine similarity in the following experimental section. The flow chart of our approach is shown in Fig. 14.

4 Results and discussion

In this section, we implement our approach of MMI for image retrieval on Corel image dataset [38–40], Brodatz texture image dataset [20, 41, 42], and WIPO global brand dataset [43, 44], which are three most widely adopted benchmark datasets in the literatures of CBIR and trademark search. In our experiments, Corel image database composed of 1000 images, named Corel-1000, is adopted to testify the effectiveness of our approach. Corel-1000 is divided into 10 categories, i.e., human beings, landscapes, departments, busses, dinosaurs, elephants, flowers, horses, mountains, and foods, and each category contains 100 images. Moreover, in order to implement rotated image retrieval, each of these images are processed through 90°, 180°, and 270°clockwise rotation. We can see from Table 5 that total 4000 images are finally generated as our dataset. Brodatz texture image dataset comprises 990 images, and they are divided into 110 categories. That is, each category contains 9 images. WIPO global brand dataset is a comprehensive source of data on the intellectual property. It contains tens of thousands of the brand data from multiple national and international sources, and has been widely used for empirical studies, reports, and factual information. Subsequently, when a query image (or trademark) is input, the feedback images (or trademarks) are ranked in the light of the relevant similarity.

Precision, recall, and F1-measure are three typical evaluation quotas in the field of information retrieval. In our experiments, all of them are utilized to evaluate the effectiveness of our approach. Precision is defined as the ratio

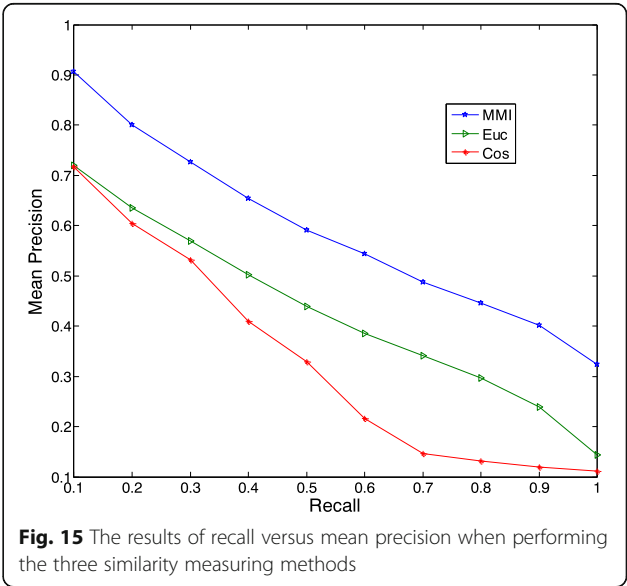


between the number of the correctly retrieved images M and the total number of the retrieved images D . Recall is defined as the ratio between M and the total number of the images A in each predefined category. Subsequently, precision P , recall R , and F1-measure are given by

$$P = M/D \tag{7}$$

Table 5 The description of Corel-1000 and Corel-4000 image datasets

	Corel-1000	Corel-4000
human beings	100	100 × 4
landscapes	100	100 × 4
departments	100	100 × 4
busses	100	100 × 4
dinosaurs	100	100 × 4
elephants	100	100 × 4
flowers	100	100 × 4
horses	100	100 × 4
mountains	100	100 × 4
foods	100	100 × 4
# total images	1000	4000



$$R = M/A \quad (8)$$

$$\text{F1-measure} = \frac{2 \times P \times R}{P + R} \quad (9)$$

In our experiments, 40 images are randomly chosen from each category in the Corel-4000 dataset query images. With regard to each querying image, we calculate the recall and the corresponding precision first. Subsequently, the recall and mean precision pair with respect to these 400 randomly chosen querying images is obtained. Figure 15 displays the experimental results comparing our similarity measuring method with the Euclidean distance and Cosine similarity methods.

We can see from Fig. 15 that our method performs much better than the Euclidean distance and Cosine similarity methods. Note that, although these two methods are widely used for measuring similarity, our method is more appropriate and achieves a better performance. That is because formula (2) normalizes the distance between MMs, and the effective denominator is taken into account in the formula.

To reveal the performance of our method of MMI, we compare it with state-of-the-art image retrieval approaches including SED [34], HSVH [29], CCM [30], DBPSP [32], MCMCM [31], and MCMCM&DBPSP [33]. In the following experiments, the pair of mean precision and recall in Fig. 16 and the mean precision with respect to first- N percentage in Fig. 17 are both exerted to demonstrate the effectiveness of MMI in depth.

From Fig. 16, we can see that for each algorithm the mean precision decreases with the raising of recall. Our approach of MMI achieves the best performance in comparison with SED, HSVH, CCM, DBPSP, MCMCM,

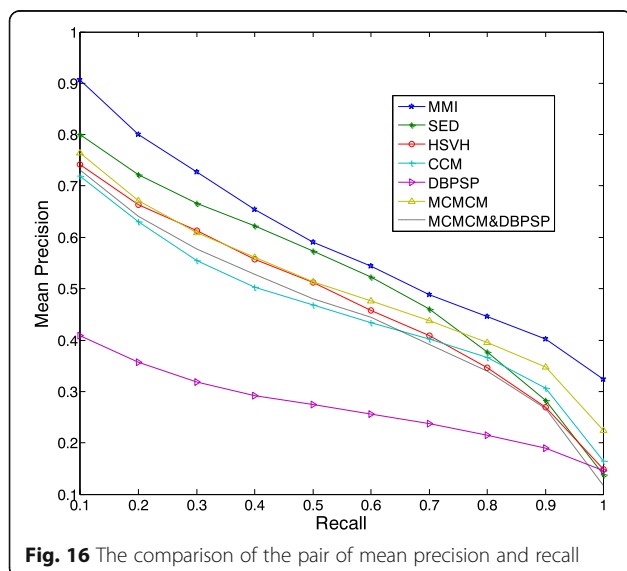


Fig. 16 The comparison of the pair of mean precision and recall

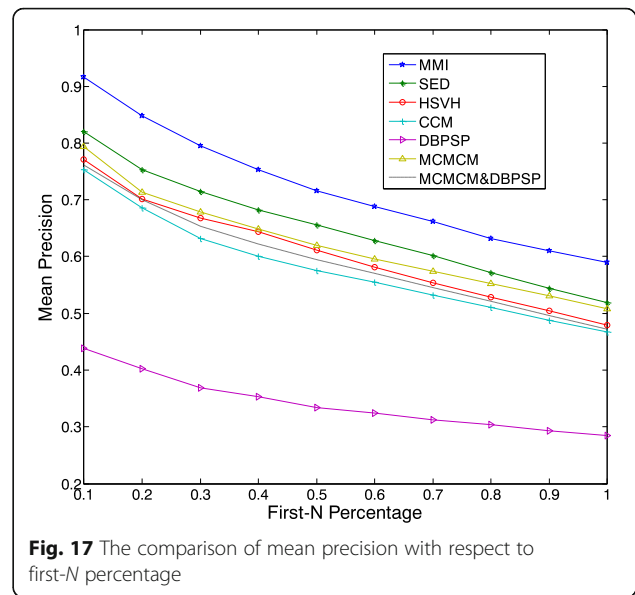


Fig. 17 The comparison of mean precision with respect to first- N percentage

and MCMCM&DBPSP. For the latter six algorithms, SED performs relatively better with mean precision of 0.800 when recall is 0.1. In contrast, the mean precision of MMI is 0.907. When recall reaches 1, MCMCM performs relatively better with mean precision of 0.224 in the latter six algorithms. By comparison, the mean precision of MMI is 0.323. For other recalls, i.e., 0.2 to 0.9, MMI achieves superior mean precisions and performs better than other six algorithms. In Fig. 17, we evaluate these seven algorithms by comparing the mean precision for first- N percentage images in each category. We can see from Fig. 17 that the experimental results clearly reveal that for the first 10–100% images of each category, MMI is significantly superior to other six algorithms. Specifically, the mean precisions of the first 10% images

Table 6 The comparison of F1-measure

	F1-measure						
	MMI	SED	HSVH	CCM	DBPSP	MCMCM	MCMCM&DBPSP
human beings	0.6208	0.5942	0.4718	0.3617	0.2334	0.4141	0.3327
landscapes	0.3135	0.2856	0.4237	0.3067	0.2270	0.4678	0.3173
departments	0.4380	0.5800	0.3963	0.3130	0.3199	0.3163	0.3513
busses	0.4191	0.5657	0.4859	0.4747	0.4524	0.4961	0.3862
dinosaurs	0.9964	0.7217	0.7468	0.8151	0.3812	0.8968	0.7862
elephants	0.5031	0.3060	0.3562	0.2648	0.2391	0.2980	0.2811
flowers	0.7746	0.7232	0.3524	0.7623	0.7138	0.6702	0.6824
horses	0.4755	0.4601	0.6953	0.3761	0.2368	0.5202	0.4714
mountains	0.2139	0.3128	0.3172	0.2817	0.1871	0.2616	0.2325
foods	0.4007	0.4276	0.3743	0.3187	0.2509	0.3084	0.4703
Mean F1-measure	0.5156	0.4977	0.4620	0.4275	0.3242	0.4650	0.4311

Table 7 The comparison of mean precision in Brodatz dataset

	6 different algorithms					
	MMI	SED	CCM	DBPSP	MCMCM	MCMCM&DBPSP
Mean precision	0.8199	0.6385	0.6546	0.3781	0.5502	0.6872

for MMI, SED, HSVH, CCM, DBPSP, MCMCM, and MCMCM&DBPSP are 0.916, 0.821, 0.771, 0.753, 0.439, 0.794, and 0.761 respectively, and when first- N percentage reaches 100%, the corresponding mean precisions are 0.589, 0.519, 0.480, 0.467, 0.285, 0.509, and 0.472 respectively.

Moreover, in view of the value of recall ranging from 0.1 to 1.0, we count the corresponding mean precision and calculate F1-measure with respect to each category, which is subsequently shown in Table 6.

We can see from Table 6 that our approach of MMI achieves the best F1-measure in four categories including human beings, dinosaurs, elephants, and flowers. Although for other categories MMI could not obtain the best performance, it still performs better than the other six algorithms in terms of the evaluation of mean F1-measure shown in the last row of Table 6.

In order to further reveal the superiority of our proposed method, the standard Brodatz dataset is applied in subsequent experiment. In specific, we select querying image from each category of Brodatz dataset. The corresponding precision is then calculated

with the recall of 1. That is, we count the precision when nine corresponding images are all retrieved with respect to the relevant category. In Table 7, our approach of MMI is compared with SED, CCM, DBPSP, MCMCM, and MCMCM&DBPSP via the mean precision.

Note that, HSVH [29] is based on HSV color space, and the images in Brodatz dataset are all png format of gray level, which cannot be converted to HSV color space. In view of such reason, the proposed MMI approach is only compared with SED, CCM, DBPSP, MCMCM, and MCMCM&DBPSP. We can see from Table 7 that our approach of MMI achieves the best retrieval performance in terms of the highest mean precision.

In order to directly display the retrieval results and reveal the performance of the proposed method, in Corel-4000 dataset, we randomly select an image from each category as a query. The similarity is then computed between the query image and the image in database. The returned images are ranked in terms of the descending order of the similarity. The image retrievals of eight examples are displayed from Figs. 18, 19, 20, 21, 22, 23, 24, and 25.

From Figs. 18, 19, 20, 21, 22, 23, 24, and 25, we can see that eight examples of image retrieval on Corel-4000 dataset are tested. For each query, the top 48 images are ranked based on their corresponding similarities. Among the returned 48 images of each query, the first four feedback images are the

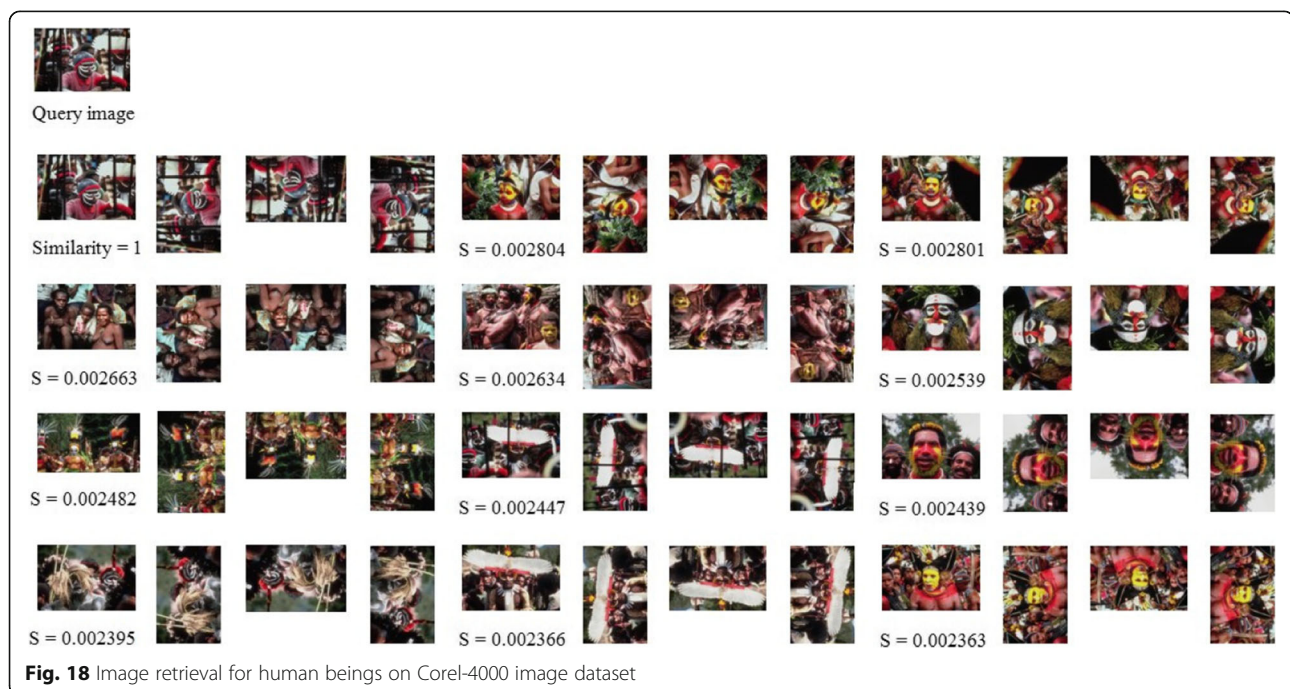


Fig. 18 Image retrieval for human beings on Corel-4000 image dataset



Fig. 19 Image retrieval for departments on Corel-4000 image dataset

original image and its three rotated images with the corresponding similarity of 1. Note that each series of four images shares the same similarity. That is because although each image is rotated by 90°, 180°, and 270° clockwise, the achieved MMs comprised of

color and texture features remain the same. Moreover, although the analogous texture and color features may disperse in different position of the returned images, these similar images can be also obtained, owing to the concept that the generated



Fig. 20 Image retrieval for dinosaurs on Corel-4000 image dataset

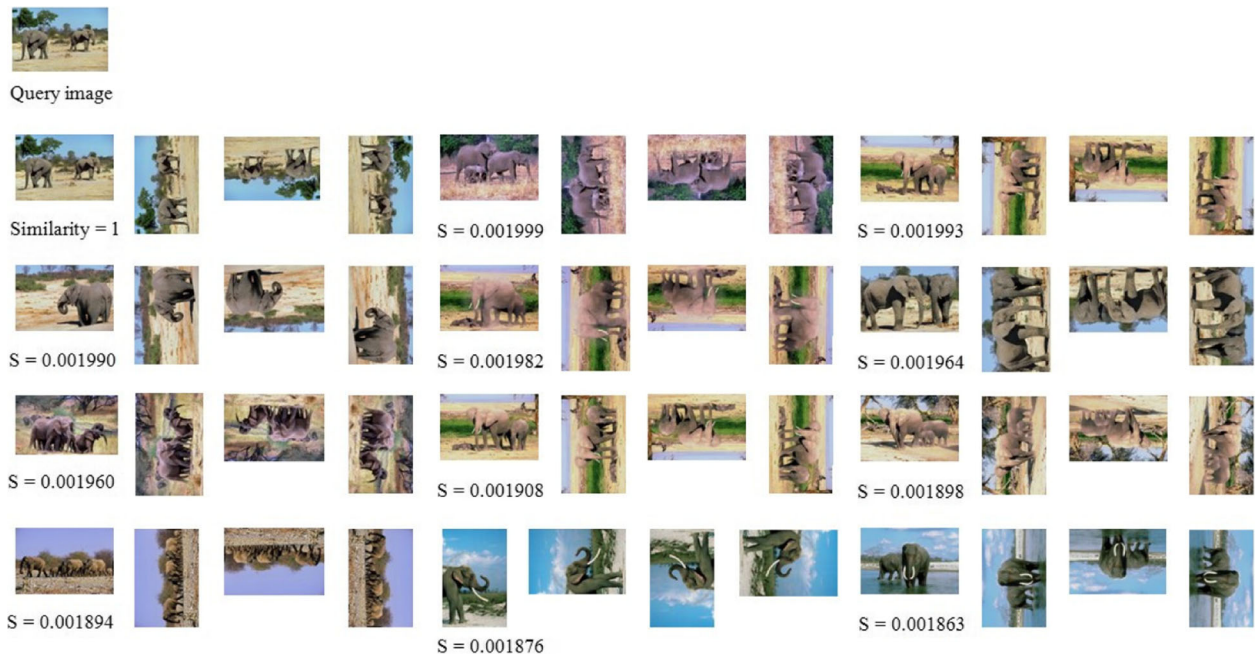


Fig. 21 Image retrieval for elephants on Corel-4000 image dataset

MMs count the times where each kind of motif shares the same average gray. Thus, such kind of retrieval strategy is much closer to the human understanding for content-based image retrieval. From the results, the retrieval precisions of human beings, departments, dinosaurs, elephants, flowers, horses, and

food categories are better than that of a landscape category. In specific, the top 48 retrieval images with regard to the first 7 queries are entirely relevant. However, for the eighth query of landscapes, some inaccurate images, e.g., series 2 about mountains and series 7 and series 11 about departments, are returned. By further analyzing such



Fig. 22 Image retrieval for flowers on Corel-4000 image dataset



Fig. 23 Image retrieval for horses on Corel-4000 image dataset

results, we find that although they come from different categories, their color and texture are indeed similar to the query image, due to the similar MMs.

In the last experiment, for the purpose of simulating trademark retrieval, the proposed method is

implemented on WIPO global brand dataset for further revealing its performance. Six typical trademark images are selected as query images, and the similarities are then computed between each query trademark and the trademarks in database. The returned

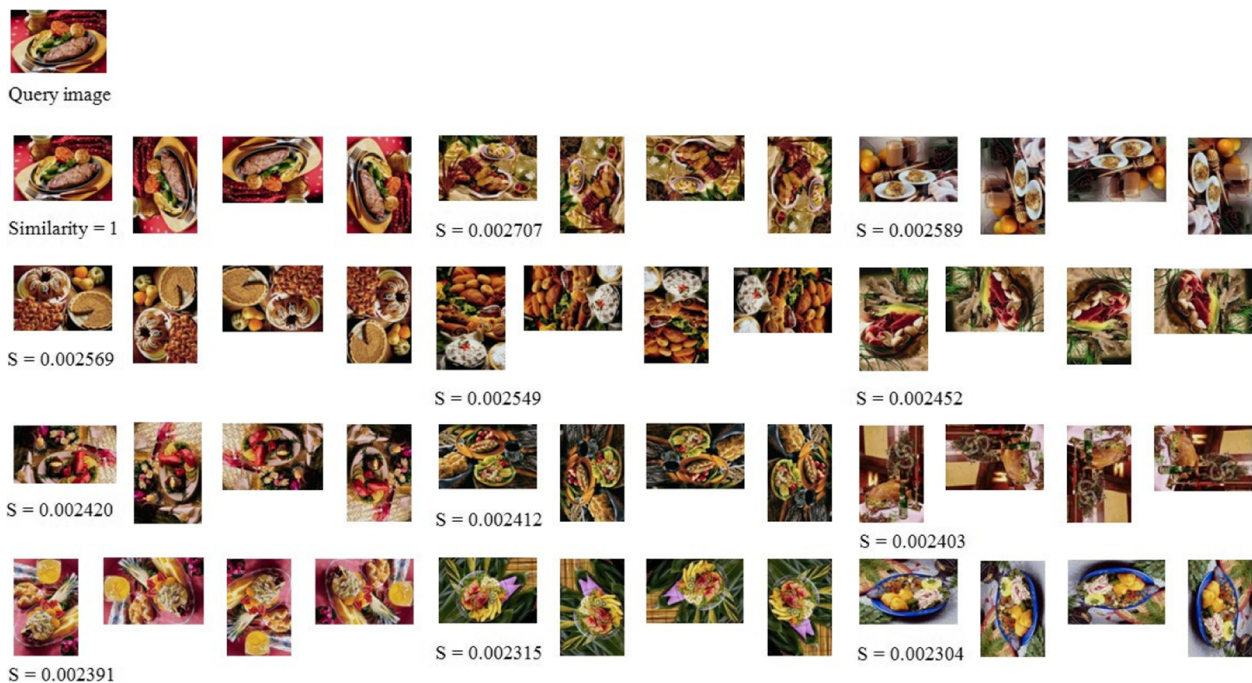


Fig. 24 Image retrieval for foods on Corel-4000 image dataset

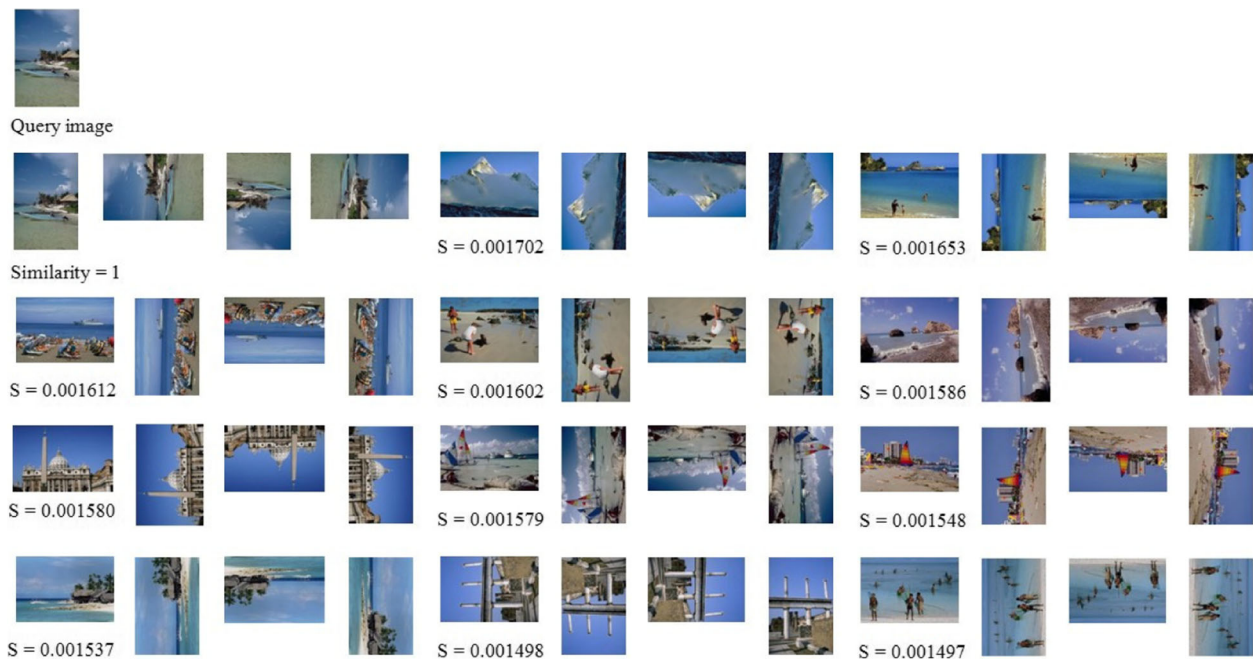


Fig. 25 Image retrieval for landscapes on Corel-4000 image dataset

trademarks are ranked in terms of the descending order of the similarity. The six examples of trademark retrieval are displayed in Fig. 26.

We can see from Fig. 26 that the six examples of trademark retrieval on WIPO global brand dataset are performed. For each query trademark, the top nine returned trademarks are ranked in terms of their similarities, and these nine trademarks are all vision relevant to the respective query. Moreover, although some partial shapes are rotated comparing to the original shapes in the query trademark, i.e., the fourth to ninth returned trademarks in Fig. 26a, the third returned trademarks in Fig. 26e, and the first returned trademarks in Fig. 26f, they are also retrieved by the proposed method. In brief, the proposed method can find the similar trademarks to its query. Moreover, the similar shapes after rotation can also be retrieved. That is to say, our method can be used as an effective tool to implement trademark retrieval, which would protect the legitimate rights and interests of registered trademarks.

5 Conclusions

In this study, eight kinds of novel motifs are first proposed to describe all textures in each 2×2 grids of an image. Subsequently, two motif-based matrices, i.e., MAM and MEM, are constructed to describe the color and texture features of an image respectively. What is more, in terms of the inference, we integrate

the advantages of both structural and statistical methods. That is, MAM and MEM are further mapped to MM to resolve the issues of rotated image retrieval. In view of MM, such a 256×8 matrix incorporates the colorful and textural characters and depicts the consistent feature between the original and its rotated images. In order to effectively measure the similarity between images, a normalized evaluation measure is utilized to calculate the similarity between the template MM and query MM. We first carry out the experiments on the benchmark Corel image dataset, and the experimental results show that our normalized evaluation measure performs better than traditionally used Euclidean distance and Cosine similarity. Subsequently, the proposed MMI is compared with SED, HSVH, CCM, DBPSP, MCMCM, and MCMCM&DBPSP on Corel image dataset and compared with SED, CCM, DBPSP, MCMCM, and MCMCM&DBPSP on Brodatz texture image dataset (As HSVH is based on HSV color space, it cannot be performed on Brodatz dataset, which are all png format of gray-level images), and the experimental results demonstrate the superiority of our method. In order to directly display the retrieval results, eight examples on Corel dataset and six examples on WIPO global brand dataset are tested, and the experimental results demonstrate the effectiveness of our method for CBIR and trademark retrieval.

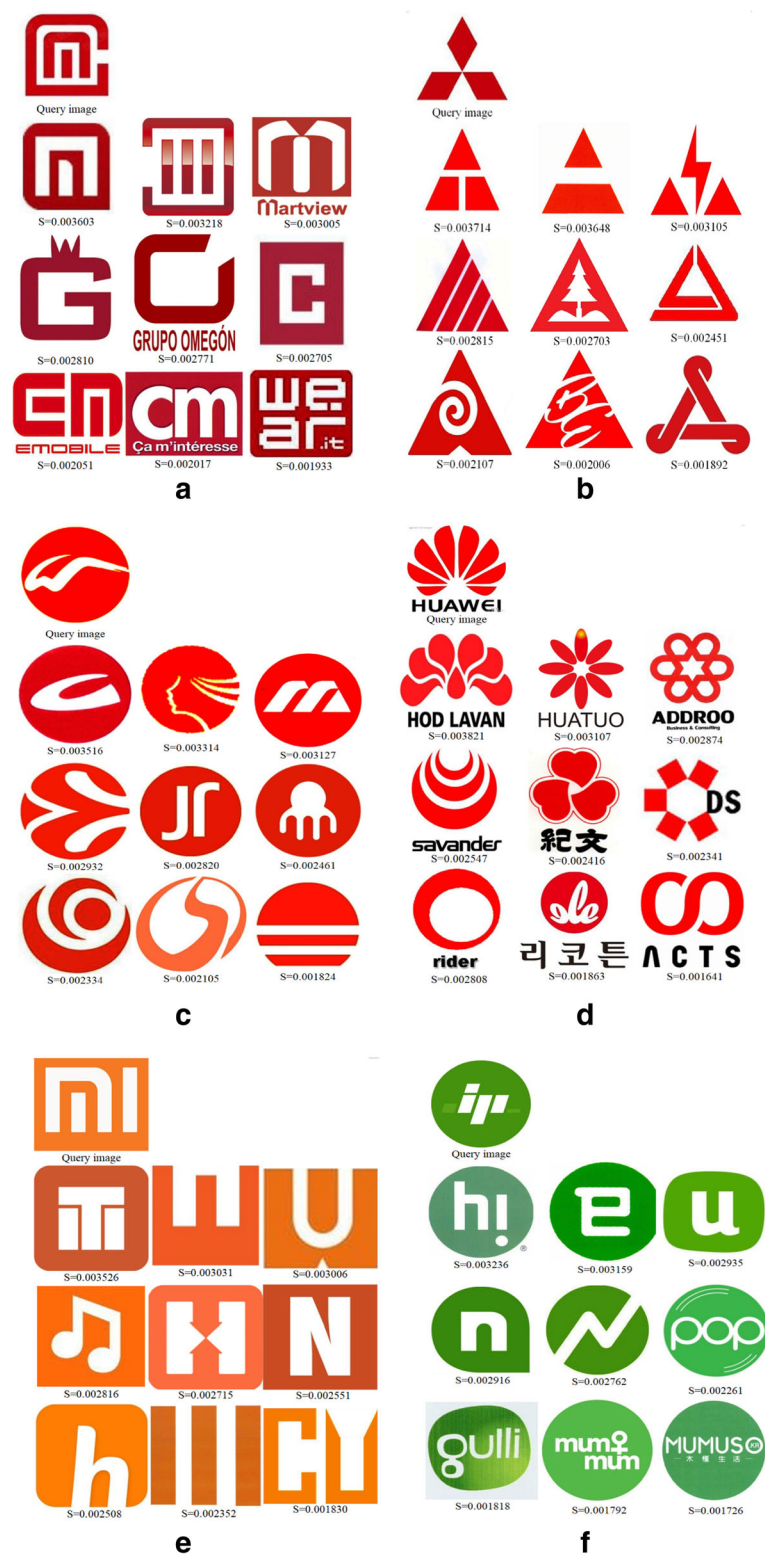


Fig. 26 The six examples of trademark retrieval. **a** The trademark retrieval for query image "m". **b** The trademark retrieval for query image "mitsubishi". **c** The trademark retrieval for query image "w". **d** The trademark retrieval for query image "HUAWEI". **e** The trademark retrieval for query image "MI". **f** The trademark retrieval for query image "ip"

Abbreviations

AGO: Ascending gray order; CBIR: Content-based image retrieval; CCM: Color co-occurrence matrix; DBPSP: Difference between pixels of scan pattern; HSV: Hue, saturation, value; HSVH: HSV histogram; MCMCM: Modified color motif co-occurrence matrix; MEM: Motif excessive matrix; MM: Motif matrix; MMI: Motif matrix inference; MULT: Minimal upper left triangle; SED: Structure elements' descriptor

Acknowledgements

The authors thank the editors and reviewers for their considerations to the publication of this paper.

Funding

This work was sponsored by National Natural Science Foundation of China (61673193), Fundamental Research Funds for the Central Universities (JUSRP51635B, JUSRP51510), China Postdoctoral Science Foundation (2017M621625), and Natural Science Foundation of Jiangsu Province (BK20150159).

Availability of data and materials

Corel image dataset at <http://wang.ist.psu.edu/docs/related/>. Brodatz texture image dataset at <http://sipi.usc.edu/database/database.php?volume=textures>. WIPO global brand dataset at <http://www.wipo.int/branddb/en/>

Authors' contributions

YX and WS contributed to the design of the motif average matrix (MAM), the motif excessive matrix (MEM), the motif matrix (MM), and drafted the manuscript. HZ, YY, and TC contributed to the design of the experiments, monitored experimental epidemics, and participated in data analysis. All the authors gave their final approval for publication.

Competing interests

The authors declare that they have no competing interests.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Author details

¹School of Internet of Things (IOT) Engineering, Jiangnan University, Wuxi, China. ²Engineering Research Center of Internet of Things Applied Technology, Ministry of Education, Wuxi, China. ³Jiangsu Provincial Engineering Laboratory of Pattern Recognition and Computational Intelligence, Jiangnan University, Wuxi, China.

Received: 6 February 2018 Accepted: 2 August 2018

Published online: 01 October 2018

References

1. B Julesz, Textons, the elements of texture perception and their interactions. *Nature* **290**(5802), 91–97 (1981)
2. B Julesz, Texton gradients: The texton theory revisited. *Biol. Cybern.* **54**, 245–251 (1986)
3. AWM Smeulders et al., Content-based image retrieval at the end of the early years. *IEEE Trans. Pattern Anal. Mach. Intell.* **22**, 1349–1379 (2000)
4. Y Liu, D Zhang, G Lu, WY Ma, A survey of content-based image retrieval with high-level semantics. *Pattern Recogn.* **40**(11), 262–282 (2007)
5. G Michele, Z Bertrand, Body color sets: A compact and reliable representation of images. *J. Vis. Commun. Image Represent.* **22**(1), 48–60 (2011)
6. W Song, Y Zhang, F Liu, Z Chai, SC Park, Taking advantage of multi-regions-based diagonal texture structure descriptor for image retrieval. *Expert Syst. Appl.* **96**(15), 347–357 (2018)
7. A El-ghazal, O Basir, S Belkasim, Invariant curvature-based Fourier shape descriptors. *J. Vis. Commun. Image Represent.* **23**(4), 622–633 (2012)
8. F Ding, YJ Wang, JY Dai, QS Li, QJ Chen, A recursive least squares parameter estimation algorithm for output nonlinear autoregressive systems using the input-output data filtering. *J. Franklin Inst.* **354**(15), 6938–6955 (2017)
9. ZH Zhang, WH Li, B Li, in *International Conference on Information Assurance and Security*. An improving technique of color histogram in segmentation-based image retrieval (Xian, IEEE, 2009), pp. 381–384
10. SH Wan, PQ Jin, LH Yue, in *International Conference on Image and Graphics*. An Effective Image Retrieval Technique Based on Color Perception (Hefei, IEEE, 2011), pp. 1017–1022
11. CC Chang, WC Wu, YC Hu, in *International Conference on Future Generation Communication and Networking Symposia*. Content-based color image retrieval system using color difference features (Sanya, IEEE, 2008), pp. 181–184
12. YY Wu, YQ Wu, in *International Conference on Image and Signal Processing*. Shape-based image retrieval using combining global and local shape features (Tianjin, IEEE, 2009), pp. 1–5
13. A Adnan, S Gul, M Ali, AH Dar, in *International Conference on International Conference on Emerging Technologies*. Content based image retrieval using geometrical-shape of objects in image (Islamabad, IEEE, 2007), pp. 222–225
14. XJ Qin, YH Yang, in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Similarity measure and learning with gray level aura matrices (GLAM) for texture image retrieval (Washington, IEEE, 2004), pp. 1–326
15. A Mosleh, Z Farzad, A Reza, in *International Symposium on Signals, Circuits and Systems*. Texture Image Retrieval Using Contourlet Transform (Lasi, IEEE, 2009), pp. 1–4
16. BB Baharum, K Ullah, in *Frontiers of Information Technology*. Efficient Image Retrieval Based on Quantized Histogram Texture Features in DCT Domain (Islamabad, IEEE, 2011), pp. 89–94
17. XF Ding, H Jin, Efficient and progressive algorithms for distributed skyline queries over uncertain data. *IEEE Trans. Knowl. Data Eng.* **24**(8), 1448–1462 (2012)
18. XF Ding, X Lian, L Chen, H Jin, Continuous monitoring of skylines over uncertain data streams. *Inf. Sci.* **184**, 196–214 (2012)
19. G Quellec, M Lamard, G Cazuguel, B Cochener, C Roux, Adaptive nonseparable wavelet transform via lifting and its application to content-based image retrieval. *IEEE Trans. Image Process.* **19**(1), 25–35 (2010)
20. S Murala, RP Maheshwari, R Balasubramanian, Local tetra patterns: A new feature descriptor for content-based image retrieval. *IEEE Trans. Image Process.* **21**(5), 2874–2886 (2012)
21. XF He, Laplacian regularized D-optimal design for active learning and its application to image retrieval. *IEEE Trans. Image Process.* **19**(1), 254–263 (2010)
22. M Flickner, H Sawhney, W Niblack, J Ashley, Q Huang, B Dom, Query by image and video content: The QBIC system. *Computer* **28**(9), 23–32 (1995)
23. A Pentland, RW Picard, S Sclaroff, Photobook: Content-based manipulation of image databases. *Int. J. Comput. Vis.* **18**(3), 233–254 (1996)
24. JR Smith, SF Chang, in *Proceedings of the 4th ACM international conference on multimedia*. VisualSEEK: a fully automated content-based image query system (Boston, ACM, 1997), pp. 87–98
25. WY Ma, BS Manjunath, Netra: A toolbox for navigating large image databases. *Multimedia Systems* **7**(3), 184–198 (1999)
26. T Gevers, AW Smeulders, Pictoseek: Combining color and shape invariant features for image retrieval. *IEEE Trans. Image Process.* **9**(1), 102–119 (2000)
27. JZ Wang, J Li, G Wiederhold, SIMPLcity: Semantics-sensitive integrated matching for picture libraries. *IEEE Trans. Pattern Anal. Mach. Intell.* **23**(9), 947–963 (2001)
28. C Carson, S Belongie, H Greenspan, J Malik, Blobworld: Image segmentation using expectation-maximization and its application to image querying. *IEEE Trans. Pattern Anal. Mach. Intell.* **24**(8), 1026–1038 (2002)
29. CH Su, HS Chiu, TM Hsieh, in *International Conference on Electrical and Control Engineering*. An efficient image retrieval based on HSV color space (Sichuang, IEEE, 2011), pp. 5746–5749
30. CH Lin, RT Chen, YK Chan, A smart content-based image retrieval system based on color and texture feature. *Image Vis. Comput.* **27**(6), 658–665 (2009)
31. M Subrahmanyam, QM Jonathan Wu, RP Maheshwari, R Balasubramanian, Modified color motif co-occurrence matrix for image indexing and retrieval. *Comput. Electrical Eng.* **39**(3), 762–774 (2013)
32. SS Yu, SY Huang, YH Pan, HC Wu, in *International Conference on Computer Symposium*. An easy dominant color extraction and edge valley histogram for image retrieval (Taiwan, IEEE, 2010), pp. 159–164
33. CH Lina, CC Chenb, HL Leeb, JR Liaob, Fast K-means algorithm based on a level histogram for image retrieval. *Expert Syst. Appl.* **41**(7), 3276–3283 (2014)
34. XY Wang, ZY Wang, A novel method for image retrieval based on structure elements' descriptor. *J. Vis. Commun. Image Represent.* **24**(1), 63–74 (2013)
35. L Vincent, Morphological grayscale reconstruction in image analysis: Applications and efficient algorithms. *IEEE Trans. Image Process.* **2**(2), 176–201 (1993)

36. A Bleau, LJ Leon, Watershed-based segmentation and region merging. *Comput. Vis. Image Underst.* **77**(3), 317–370 (2000)
37. JB Kim, HJ Kim, Multiresolution-based watersheds for efficient image segmentation. *Pattern Recognit. Lett.* **24**(1), 473–488 (2003)
38. D Xu, S Yan, D Tao, S Lin, HJ Zhang, Marginal fisher analysis and its variants for human gait recognition and content-based image retrieval. *IEEE Trans. Image Process.* **16**(11), 2811–2821 (2007)
39. L Yang, R Jin, L Mummert, R Sukthankar, A Goode, B Zheng, M Satyanarayanan, A boosting framework for visuality-preserving distance metric learning and its application to medical image retrieval. *Pattern Anal. Mach. Intell.* **32**(1), 30–44 (2010)
40. XH Yang, LJ Cai, Adaptive region matching for region-based image retrieval by constructing region importance index. *Comput. Vis.* **8**(2), 141–151 (2014)
41. Y Ji, F Ding, Multiperiodicity and exponential attractivity of neural networks with mixed delays. *Circuits Syst. Signal Process.* **36**(6), 2558–2573 (2017)
42. X Li, Y Ma, W Yu, Geometry-invariant texture retrieval using a dual-output pulse-coupled neural network. *Neural Comput.* **24**(1), 194–216 (2012)
43. R Setchi, FM Anuar, Multi-faceted assessment of trademark similarity. *Expert Syst. Appl.* **65**(1), 16–27 (2016)
44. X Li, DQ Zhu, An improved SOM neural network method to adaptive leader-follower formation control of AUVs. *IEEE Trans. Industrial Electronics.* **65**(10), 8260–8270 (2018)

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)