Color-Texture Based Two-Layer Image Retrieval

Yung-Kuan Chan, Chi-Shiang Chan, Tsung-Hsin Chen, and Kuan-Cheng Lin

Abstract—in this thesis, an original descriptor for image retrieval improves by combine two different features. The original descriptor is Micro-Structure Descriptor, which denotes color and texture of images. The other feature is extracted by gray-level co-occurrence matrix. The Particle Swarm Optimization also applies to improve the performance. The experiment result shows that this method is than original Micro-Structure Descriptor.

Keywords—Content-based image retrieval (CBIR), micro-structure descriptor (MSD), Particle Swarm Optimization (PSO)

I. INTRODUCTION

THIS paper introduces a new descriptor for Content-based image retrieval. CBIR has developed for many years so far, the features of images can be easily categorized to three main domains: color [1], shape [2], and texture [3, 4]. Different features can be retrieved from different types of images.

This paper focuses on natural database which contains some natural objects. For example, Corel 1k contains fruits, sea, horses…etc. Micro-structure descriptor (MSD) [5] is one of CBIR which uses natural database. MSD combines texture features with color features by detecting the micro-structure and recording the colors within the micro-structure. The color feature histogram is constructed by the quantity of accumulation, and MSD is a feature vector which records this quantity of accumulation.

Since texture features of MSD have not yet completely developed, discrimination power of MSD can increase. Extracting features from blocks creates a lack of global feature to search within MSD. MSD is a 72-length feature vector, but not each value denotes the feature well. This paper descending MSD dimensions first. Decreasing the length of feature vector makes the descriptor consumes less storage. This paper descending to search within MSD. MSD is a 72-length feature vector, but not each value denotes the feature well. This paper descending MSD dimensions first. Decreasing the length of feature vector makes the descriptor consumes less storage. This paper descending to search within MSD.

This paper not only constructs new descriptor but also changes the traditional methods of retrieval. The MSD retrieval method returns the most similar 12 images of the query image. The fixed return number let precision is limited. This paper tries to change the return images. It intends to divides the existing database to some different clusters. The retrieval will return all similar pictures in the cluster related to the query image.

However, some clusters may be too large to get correct results. As a result, the larger clusters need to be divided into some smaller clusters. Particle swarm Optimization (PSO) is an unsupervised algorithm to train parameters. PSO is used to train the best cluster number should be in this paper. PSO is also used to improve the new descriptor by applying weight into features.

By constructing the new feature vector and changing the retrieval method, this paper provides a better descriptor for image retrieval. This descriptor is smaller than MSD so it consumes less storage.

II. RELATED WORK

A. HSV color space

HSV color space is a cylindrical-coordinate color space. The H component means hue, which denotes color type. It is designed as a cycle, 0° is red, 120° is green, and 240° is blue. The S component means saturation, which denotes color purity. The V component means value, which denotes brightness. HSV color space is often used in image analysis.

B. Sobel operator

Edge detection is an image processing technique used in different applications. In image analysis, Sobel operator [8, 9] is a popular edge detection method. Sobel operator calculates the gradient of a pixel by referring its around pixels. More precisely, Sobel operator employs two 3×3 masks as Fig.1 (a), (b) the left one is for vertical edge and the right one is for horizontal edge to compute the gradient in the image.

\[
\begin{array}{ccc}
-1 & 0 & +1 \\
-2 & 0 & +2 \\
-1 & 0 & +1 \\
\end{array}
\]

\[
\begin{array}{ccc}
+1 & +2 & +1 \\
0 & 0 & 0 \\
-1 & -2 & -1 \\
\end{array}
\]

(a) Vertical Mask (b) Horizontal Mask

(c) 3×3 Block

Fig. 1 Masks of Sobel

Sobel operator calculates gradient of the block as follow equation:

\[
G_x = (p_1 + 2p_6 + p_9) - (p_1 + 2p_4 + p_7) \\
G_y = (p_1 + 2p_2 + p_3) - (p_7 + 2p_6 + p_9)
\]

where \( p_i, i=1,2,3,\ldots,9 \) are the value of each location in a block in an image as shown in Fig. 1(d). After gradient is computed, the gradient direction can also be computed as follow equation:

\[
\theta = \tan \frac{G_y}{G_x}
\]

Chi-Shiang Chan is with the Asia University, Taiwan, ROC, (e-mail: CSChan@asia.edu.tw).

Yung-Kuan Chan, Tsung-Hsin Chen, and Kuan-Cheng Lin are with the National Chung Hsing University, Taiwan, ROC, (phone: 886-2-22840864; e-mail: ykchan@nchu.edu.tw, sundogsii@gmail.com, kclin@nchu.edu.tw).

[1] Yung-Kuan Chan, [2] Tsung-Hsin Chen, and [3] Kuan-Cheng Lin are with the National Chung Hsing University, Taiwan, ROC, (e-mail: ykchan@nchu.edu.tw, sundogsii@gmail.com, kclin@nchu.edu.tw).
The gradient direction is assistance for image analyze.

C. Micro-structure descriptor (MSD)

General CBIR extracts three kinds of features, including color, shape, and texture. In MSD, color and texture are considered as features, and two features are mixed up to get color histogram. Therefore the method in this thesis uses HSV color model to calculate features.

HSV is cylindrical-coordinate color model, which rearrange vectors are calculated by four steps. Feature vectors can be constructed. More precisely, the feature vectors are calculated by four steps:

1. After calculating the gradient direction \( \theta \) of all pixels, they are marked.
2. Sobel Operator is performed to extract horizontal and vertical orientations of a pixel. That is, the horizontal and vertical orientations of each color pixel. Therefore, after marking those pixels, the final result is shown in Fig. 2(d).
3. Construct feature vectors by accumulates their histogram. The feature vector value is calculated by equation below:
   \[
   \text{H}(w_0) = \{N[f(p_i) = w_0]\wedge f(p_i) = w_i | p_i - p_0| = 1\}
   \]
   where \( w_0 = w_i, i \in \{1, 2, \ldots, B\} \)

D. Gray level co-occurrence matrix (GLCM)

Gray level co-occurrence matrix denotes the pixel-value and the other pixel value. It will construct a square matrix by the equation below:
\[
M_{ij} = \sum_{k=1}^{m} \sum_{l=1}^{n} \begin{cases} 
1 & \text{if } k_{ij} = i \text{ and } l_{\text{neighbor}} \\
0 & \text{otherwise}
\end{cases}
\]

Where \( M_{ij} \) represents the value at \( i \) row and \( j \) column in co-occurrence matrix, \( m \) is the wide and \( n \) is the high, \( k_{ij} \) is \( k \) row and \( l \) column of image, \( l_{\text{neighbor}} \) has different direction as below:

After calculating the gradient direction \( \theta \) of all pixels, they are quantized into \( m \) bins for defining micro-structure.

A micro-structure, utilizes \( 3 \times 3 \) block as a unit to find the positions with similar attributes. More precisely, if the neighboring pixels of a center pixel have the same values with center pixel in the \( 3 \times 3 \) block, these pixels are marked. Take an example in Fig. 2, the value of the center pixel is five. After checking the eight neighboring pixels, three pixels have the same pixel value with the center pixel. Therefore, after marking these pixels, the final result is shown in Fig. 2(d).

Fig. 2 The way to obtain micro-structure (a) one of \( 3 \times 3 \) block of edge orientation constructed by STEP2 (b) compare center pixel with 8-neighbors (c) keeps the same values (d) the micro-structure
After co-occurrence matrix has built. There are some equations for extracting some features as below:

\[
\begin{align*}
\text{Uniformity (UNI)} &= \sum C_{ij}^2 \\
\text{Entropy (ENT)} &= \sum C_{ij} \log C_{ij}
\end{align*}
\]

These features are often used to be texture features.

E. K-means

Cluster means those data in the same cluster has similar features. Clustering is an algorithm for find cluster in a data set. Cluster technology is researched for many years and it is often classified into two major methods: Hierarchical and Partitioning [6]. Hierarchical uses some hierarchical structure to decomposition data into different class. Partitioning clustering divides data set into designed number of distinct clusters. K-means is a kind of partitioning clustering which is used widely. K-means is a simple unsupervised learning algorithm that means computer aids people to cluster the data. The main idea of K-means is minimized the error for number of clusters. The error function is as below:

\[
E = \sum_{i=1}^{K} \sum_{x \in S_k} ||x - C_k||^2
\]

Where K is the number of clusters to divide, \(C_k\) is the \(k\)th center of cluster, \(x\) denotes data belong to cluster \(S_k\). The K-means is done by steps below:

**STEP1. Random cluster center**

If we want to divide data in \(k\) clusters, we should random \(k\) cluster centers. Take two dimension for example, let these clusters separate into the space.

**STEP2. Cluster**

Distinguish data bellows which cluster by the distance with \(k\) cluster centers, so there are \(k\) cluster data.

**STEP3. Change cluster center**

Change position of cluster center by calculating each cluster mean. Let each cluster mean become new cluster center.

\[
M = \frac{1}{n} \sum_{x \in S_k} x
\]

**STEP4. Repeat STEP3.**

Renew the cluster centers by Step3 until cluster centers do not change.

F. Particle Swarm Optimization (PSO)

PSO [7] is used to find the optimization parameter of other algorithm. It is an optimization algorithm to simulate actions of birds block. PSO regards the parametric combination as a particle. This particle finds the best position to make best results. It will be introduced step by step as follow:

**STEP1. Define the fitness function**

Define the fitness function and the performance value to check the particle. This function has \(m\) input value and one output value.

**STEP2. Initialization**

Initial the PSO with random \(n\) locations of particles and \(n\) velocities. \(n\) is the number of particles in a swarm.

**STEP3. Parameter evaluate**

Evaluate performance of each particle. The algorithm records the particle which has best performance as particle’s best (pbest).

**STEP4. Generate new particle**

Generate new particle by pbest and gbest as follow equation:

\[
v_{i, t+1} = w \times v_{i, t} + c_1 \times r_1 \times (pbest_i - s_{i, t}) + c_2 \times r_2 \times (gbest - s_{i, t})
\]

\[
s_{i, t+1} = s_{i, t} + v_{i, t+1}
\]

Where \(v_{i, t}\) is the velocity of particle \(i\) at iteration \(k\), \(w\) is the weight decrease the velocity, \(c_1\) are coefficients with value of 2, \(s_{i, t}\) is the current position of particle \(i\) at iteration \(k\), pbest\(_{i}\) is the pbest of particle \(i\).

**STEP5. Repeat**

Repeat Step3–4 until getting the optimization.

### III. THE PROPOSED METHOD

**A. Features**

1. Micro-structure descriptor (MSD) features

MSD is introduced in section 2.3. It describes some features from image, but no idea about whether each feature has good discrimination power. Discrimination power is an index of recognizing images. This paper calculates the discrimination power of MSD features first. In statistics, F-value is used to denote the discrimination power. F-value is easy to compute by formula below:

\[
F = \frac{\text{between-group variability}}{\text{within-group variability}}
\]

\[
\text{between-group variability} = \sum_{i} n_i (\bar{Y}_i - \bar{Y})^2 / (K - 1)
\]

\[
\text{within-group variability} = \sum_{ij} (Y_{ij} - \bar{Y}_i)^2 / (N - K)
\]

Where \(\bar{Y}_{ij}\) denotes every data, \(\bar{Y}_i\) denotes ith cluster center, \(K\) is the number of cluster, \(N\) is the all data number.

2. Gray-level Co-occurrence Matrix (GLCM) features

GLCM is widely used to describe special features of image. GLCM has many different calculate formula for different features [8]. This paper calculates their F-value, which is introduction in section 3.1, to retain better features.

**B. Feature weight**

This paper constructs a new feature vector. This feature vector has different value to denote different feature. The proposed method applies some weight in coefficient of each feature to improve the precision. Besides, the weights of feature power are also applied. The different power means different
distance calculated methods. Square root means Euclidean distance and one of the power means city-block distance. Different distance measure methods produce different results, so this method considers the power is a parameter. This paper added these parameters into PSO to get best parameter group. The similar comparison between feature vectors is calculated as below:

\[ E = \sum_{i=1}^{n} w_1 \times (X_i - Y_i)^{w_2} \]  

Where \( n \) is the feature vector length, \( w_1 \) is the coefficient weight, \( X_i \) is the \( i \)th feature in one feature vector, \( Y_i \) denotes another feature vector, \( w_2 \) is the power weight

**C. Classification**

There are some classes in natural data set. There are some data called outlier, which are not as similar as other data in the same class. The outliers may let the class become huge and affect the ability in classification. The huge cluster may let the cluster error as below:

![Fig. 5 Classification](image)

(a) superclass

(b) subclass

Fig. 5 Classification

In Fig. 5(a), the \( X_i \) should belong to \( C_2 \), but it is closer to \( C_1 \). In the classification algorithm, it will be departed to \( C_1 \). In Fig. 5(b), the super class should be parted in some subclass. Let the within-group variability become smaller to decrease the error.

**D. Particle Swarm Optimization (PSO)**

PSO is a good method to find optimization parameters to make best results. The proposed method takes coefficient weight, weight of feature power, and the subclass number as parameters.

Step 1: Generate Particle

Random the parameters as Table below:

**Table III.1 PARAMETERS Of PSO**

| \( C_1 \) | \( C_2 \) | \( \ldots \) | \( C_m \) | \( w_{11} \) | \( w_{12} \) | \( \ldots \) | \( w_{1n} \) | \( w_{21} \) | \( w_{22} \) | \( \ldots \) | \( w_{2n} \) |

where \( m \) is the class number, \( n \) is the feature vector length

Step 2: Fitness definition.

The performance of the parameters is the precision. This method trains data with training data, and get precision with test data.

**Train process:**

Set \( C_i \) as the \( k \) of K-means, which means \( i \)th class should be divided \( C_i \) cluster. The K-means is minimized error with weight as below:

\[ E = \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij} (X_j - S_i)^{w_2} \]  

Follow the steps in section 2.5 and the mean is changed as below:

\[ M = \frac{1}{n} \sum_{i=1}^{n} w_i \times x \]  

**Precision calculate:**

Take query image comparison with S cluster center. Choose the most close cluster center and recall all of the cluster images. The precision is calculated as below:

\[ P = \frac{n}{m} \sum_{i=1}^{n} \frac{n}{m} \]  

Where \( m \) is the recall images number, \( n \) is the correct recall images number, \( k \) is the test data number.

The PSO will optimization the precision.

**E. Mix descriptor**

In image retrieval, there are many different kinds of databases. Different databases have different feature, ex. color, texture, shape, and spatial...ETC. This paper is focus on natural databases. Some databases are designed for one kind of feature; however natural databases keep close to general life. If we just check one kind of feature in natural databases, some class will be determined error. So this method combines two kinds of features to make this mix descriptor has better discrimination. MSD features denote the color and some texture features of image and GLCM features denote texture and global features. In Fig. 6, there is an example for retrieval error.

![Fig. 6 similar images in MSD](image)

The sea is similar with the blue bus in MSD features. Combine with other features can prevent this situation happen. We consider MSD feature are weak at texture features. This method mixes the MSD and GLCM descriptor for strength the MSD.
IV. EXPERIMENT RESULTS

A. Data set

This paper takes corel 1k as the test data set. The data set has 10 different classes, which have different kind of natural. Each class has 100 similar object images as below:

B. Feature Selection

In section 3.1, this paper introduces the features selected. The F-value of each MSD feature is as Fig. 8 (a)

Though there are 72 values in MSD, but they are different discrimination power. In the Fig 8 (a), we can notice that the discrimination power of values is low. This paper sets the 30th MSD feature as threshold, so this paper chose 30 MSD features which have better F-value to reduce MSD feature vector length.

As the section above, the GLCM discrimination power is as Fig. 8 (b) This paper sets the 10th F-value as threshold, so this paper chooses 10 GLCM features which have better discrimination power. Combine these 10 features and 30MSD features which are introduced in section 3.1 for construct color-texture descriptor.

C. Particle Swarm Optimization

**MSD**

The PSO parameter is described in section 3.4. The m equals 10, n is 30 in MSD. The PSO iteration is as below:

Where the red line denotes global best, the blue line denotes particle best, the green line is average results. The precision arrives 0.72 in the end.

The result as below:

<table>
<thead>
<tr>
<th>C</th>
<th>5</th>
<th>5</th>
<th>2</th>
<th>4</th>
<th>4</th>
<th>4</th>
<th>4</th>
<th>4</th>
<th>1</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>w_1</td>
<td>0.82</td>
<td>0.88</td>
<td>0.75</td>
<td>0.5</td>
<td>0.18</td>
<td>0.61</td>
<td>0.37</td>
<td>0.66</td>
<td>0.65</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>509</td>
<td>786</td>
<td>464</td>
<td>656</td>
<td>2528</td>
<td>314</td>
<td>971</td>
<td>188</td>
<td>1.76</td>
<td>3.05</td>
</tr>
<tr>
<td>w_2</td>
<td>0.95</td>
<td>0.76</td>
<td>0.5</td>
<td>0.96</td>
<td>0.37</td>
<td>0.81</td>
<td>0.80</td>
<td>0.80</td>
<td>0.75</td>
<td>0.55</td>
</tr>
<tr>
<td></td>
<td>983</td>
<td>559</td>
<td>760</td>
<td>590</td>
<td>192</td>
<td>300</td>
<td>700</td>
<td>360</td>
<td>352</td>
<td>35</td>
</tr>
<tr>
<td>w_3</td>
<td>0.81</td>
<td>0.74</td>
<td>0.9</td>
<td>0.78</td>
<td>0.78</td>
<td>0.67</td>
<td>0.51</td>
<td>0.87</td>
<td>0.82</td>
<td>0.71</td>
</tr>
</tbody>
</table>

**GLCM**

The PSO parameter is described in section 3.4. The m equals 10, n is 30 in MSD. The PSO iteration is as below:

Where the red line denotes global best, the blue line denotes particle best, the green line is average results. The precision arrives 0.53 in the end.
The result as below:

<table>
<thead>
<tr>
<th>C</th>
<th>4</th>
<th>2</th>
<th>3</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>4</th>
<th>3</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>w1</td>
<td>0.98</td>
<td>0.12</td>
<td>0.84</td>
<td>0.28</td>
<td>0.22</td>
<td>0.12</td>
<td>0.57</td>
<td>0.20</td>
<td>0.73</td>
<td>0.26</td>
</tr>
<tr>
<td>w2</td>
<td>1.06</td>
<td>1.89</td>
<td>2.48</td>
<td>3.09</td>
<td>1.23</td>
<td>0.46</td>
<td>0.96</td>
<td>0.94</td>
<td>1.43</td>
<td>3.19</td>
</tr>
</tbody>
</table>

MIX

The PSO parameter is described in section 3.4. The m equals 10, n is 30 in MSD. The PSO iteration is as below:

![Fig. 11 Iteration of MIX PSO](image)

where the red line denotes global best, the blue line denotes particle best, the green line is average results. The precision arrives 0.778 in the end.

The result as below:

<table>
<thead>
<tr>
<th>C</th>
<th>5</th>
<th>2</th>
<th>2</th>
<th>2</th>
<th>1</th>
<th>5</th>
<th>5</th>
<th>5</th>
<th>1</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>W_m</td>
<td>0.559461262</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>W_g</td>
<td>0.559461262</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Where W_m is the weight of MSD, W_g is the weight of GLCM.

D. Comparison

The comparison of method is as below:

![Fig. 12 Precision of different methods](image)

In Fig. 12, the PSO MSD is better than MSD original. However, we apply GLCM to strong the descriptor again. This method gets better result by mix descriptor.

V. Conclusion

This paper has proposed a new descriptor for image retrieval. This method decreases the feature vector length by statistics. This descriptor combines MSD and GLCM features to denote the color and texture features well. Last, this paper optimizes the weight to get better performance. In doing so, the precision is better than original MSD.

REFERENCES