Abstract—This paper approaches the image retrieval system on the base of visual features local region RBIR (region-based image retrieval). First of all, the paper presents a method for extracting the interest points based on Harris-Laplace to create the feature region of the image. Next, in order to reduce the storage space and speed up query image, the paper builds the binary signature structure to describe the visual content of image. Based on the image’s binary signature, the paper builds the SG (signature graph) to classify and store image’s binary signatures. Since then, the paper builds the image retrieval algorithm on SG through the similar measure EMD (earth mover’s distance) between the image’s binary signatures. Last but not least, the paper gives an image retrieval model RBIR, experiments and assesses the image retrieval method on Corel image database over 10,000 images.

Keywords—Image Retrieval; Binary Signature; Signature Graph

I. INTRODUCTION

There are three common approaches for query image [1], including: image retrieval based on TBIR key word (text-based image retrieval), image retrieval based on CBIR content (content-based image retrieval) and image retrieval based on semantic SBIR (semantic-based image retrieval). The image retrieval through TBIR key word will be difficult and time-consuming in the description of image’s content. Thus, it is necessary to build CBIR retrieval system through the content of image to find out the similar image. Furthermore, when querying image through a key word or an index, we will not describe visual characteristics of image. So, we need to create a method of extracting image’s characteristics, thence reflect the content of image to find out images similar content. Extracting visual features of image is an important task of the image retrieval process on the base of content. However, if we query and compare directly the content of image, the problem will be complicated, time-consuming and costly storage space. For this reason, when comparing the image’s content, still ensuring the query speed and storage space.

A number of works related to the query image’s content has been published recently as: Extracting image objects based on the change of histogram value [1]. Similarity image retrieval based on the comparison of characteristic regions and the similar relationship of feature regions on image [2]. Color image retrieval based on the detection of local feature regions by Harris-Laplace [3], Color image retrieval based on bit plane and Lab color space [4], Converting color space and building hash table in order to query the content of color images [5],...

II. THE SIMILAR MEASURE

A. Image’s Binary Signature

In accordance [7], the binary signature is formed by hashing the data objects, and will have $k$ bit 1 and $(m - k)$ bit 0 in the bit chain $[1..m]$, with $m$ is the length of the binary signature [7]. The data objects and the object of the query are encoded on the same algorithm. When the bits in the signature data object are completely covered with the bits in the query signature, then this data object is a candidate fulfilling the query. According to [7] there are three cases: (1) the data object matches the query: every bit in the $s_q$ is covered with the bits in the signature $s_i$ of the data object (i.e., $s_q \land s_i = s_q$); (2) the object does not match the query (i.e., $s_q \land s_i \neq s_q$); (3) the signatures are compared and then give the false drop result.

Definition 1: Setting $F = (F_1, F_2, ..., F_{m})$ is a vector to describe the feature regions of image. Setting vector $F(R'_i) = (F_1(R'_i), F_2(R'_i), ..., F_m(R'_i))$ is a vector value of feature regions which was standardized on [0, 1] (i.e. $F_j(R'_i) \in [0,1], \ j = 1,...,n_F$). Setting $B'_i = b'_1 b'_2 ... b'_m$ with $b'_k = 1$ if $k = \lceil (F_j(R'_i) + 0.05) \times m \rceil$, otherwise $b'_k = 0$ $(k = 1,...,m)$. At that time, the binary signature of feature region $R'_i$ be defined as: $\text{Sig}(R'_i) = B'_1 B'_2 ... B'_m$. The binary signature of image $I$ will be $\text{Sig}(I) = \bigcup_i \text{Sig}(R'_i)$.

B. EMD Distance

Setting $I$ is a set of suppliers, $J$ is a set of consumers, $c_{ij}$ is the transportation cost from the supplier $i \in I$ to the consumer $j \in J$, we need to find out flows $f_{ij}$ to minimize the
total cost $\sum_{i\in I} \sum_{j\in J} c_{ij}f_{ij}$ with the constraints [8]: $f_{ij} \geq 0, \sum_{i\in I} f_{ij} \leq y_j, \sum_{j\in J} f_{ij} \leq x_i, i\in I, j\in J$. With $x_i$ is the provider’s general ability $i \in I$, $y_j$ is the total need of the consumer $j \in J$. The feasible condition is $\sum_{j\in J} y_j \leq \sum_{i\in I} x_i$. The EMD distance [8] as follow:

$$EMD((x,y)) = \frac{1}{m} \sum_{i=1}^{m} \frac{d_i}{f_i}$$

The image $I$ with the binary signature is $SIG_I = B^I_1B^I_2...B^I_n$, with $B^I_j = b^I_1b^I_2...b^I_m$, $b^I_j \in [0,1]$. The weight of $B^I_j$ component as: $w^I_j = w(B^I_j) = \frac{\sum_{i=1}^{m} (b^I_i \times \frac{1}{m} \times 100)}{m}$. Therefore, we have the weight vector of image $I$ will be $W_I = (w^I_1, w^I_2, ..., w^I_m)$. Setting $J$ is the image needs to calculate the similarity corresponding to the image $I$, so we need to minimize the cost of converting color distribution is as: $\sum_{j=1}^{n} d_{ij}f_{ij}$, with $F = (f_{ij})$ as a matrix of color distribution flows between $c^I$ and $c^J$, and $D = (d_{ij})$ as a matrix of Euclidean distance in RGB color space between $c^I$ and $c^J$. Setting $W_m = \min(\sum_{i=1}^{m} w^I_i, \sum_{i=1}^{m} w^J_i)$ and $W_m = \max(\sum_{i=1}^{m} w^I_i, \sum_{i=1}^{m} w^J_i)$, then the distance 

$$EMD(I,J) = \frac{1}{m} \sum_{i=1}^{m} \frac{d_{ij}}{f_{ij}}$$

is a color distribution flows from the image with the color weight $W_m$ to the image with color weight $W_m$. For this reason, the similar measure between two images $I$ and $J$ on the base of EMD distance will be the minimum cost:

$$EMD(I,J) = \min_{F[\mathbb{R}^m]} \frac{1}{m} \sum_{i=1}^{m} \frac{d_{ij}}{f_{ij}}$$

with $\sum_{i=1}^{m} f_{ij} = W_m = \min(\sum_{i=1}^{m} w^I_i, \sum_{i=1}^{m} w^J_i)$

III. DATA STRUCTURE AND IMAGE RETRIEVAL ALGORITHM

A. The Signature Graph

**Definition 2:** The Signature Graph $SG = (V, E)$ is the graph which describes the relationship between the image, has a set of vertices $V = \{oid, I, Sig(R^I)\}$ in $\mathbb{I}$ and a set of edges $E = \{(I, J) | \phi(I, J) = \phi(R^I, R^J), \forall I, J \in I \}$. With the weight of each edge $(I, J)$ is a measurement function of the similarity $\phi(I, J) = \phi(R^I, R^J)$, $\phi(R^I, R^J)$ is a threshold value and $\mathbb{I}$ is an image database. In this paper, the similar measure $\phi(I, J) = EMD(I, J) = EMD(Sig(I), Sig(J))$.

Each vertex $v \in V$ in $SG$ will determine $k$ element which has the nearest similar measurement. However, if the number of image in a large database, we will be difficult to determine the set of similar image corresponding to the query image. Therefore, we build the notion of S-KGraph so that every vertex include $k$ nearest image and call as $k$-neighboring image.

**Definition 3:** A cluster $V_i$ has center $I_i$, with $k_i\theta$ is a radius, be defined as follow:

$$V_i = V_i(I_i) = \{J | \phi(I_i, J) \leq k_i\theta, J \in \mathbb{I}, i = 1,...,n\} \cup _{k_i \in N'}$$

**Definition 4:** ($S-k$Graph) Giving a set $\Omega = \{V_i | i = 1,...,n\}$ is a set of clusters, with $V_i \cap V_j = \emptyset, i \neq j$. The $S-k$Graph = $(V_{SG}, E_{SG})$ is the graph which have the weight, including vertex set $V_{SG}$ and edge set $E_{SG}$ be defined as follow:

$$V_{SG} = \{\emptyset\} | \emptyset \in V, \forall I \in V, \phi(I_0, I) \leq k_0\theta, i = 1,...,n\}

with $d(V_i, V_j)$ is the weight between two clusters and $\forall I \in V, \phi(I_0, I) \leq k_0\theta$.

**Definition 5:** (The Rules of Distribution of Image) Giving set $\Omega = \{V_i | i = 1,...,n\}$ is a set of clusters, with $V_i \cap V_j = \emptyset, i \neq j$, setting $I_0$ is an image needs to distribute in a set of clusters $\Omega$, setting $I_i$ is a center of cluster $V_i$ so that $\phi(I_0, I_m) - k_0\theta = \min_{i}(\phi(I_0, I_i) - k_0\theta), i = 1,...,n$ with $I_i$ is a center of cluster $V_i$. There are three cases as follow:

1. If $\phi(I_0, I_m) - k_0\theta$ then the image $I_0$ will be distributed in cluster $V_m$.
2. If $\phi(I_0, I_m) > k_0\theta$ then setting $k_0 = [\phi(I_0, I_m) - k_0\theta]/\theta$, at that time:
   (2.1) If $k_0 > 0$ then creating cluster $V_0$ with center $I_0$ and radius is $k_0\theta$, at that time $\Omega = \Omega \cup \{V_0\}$.
   (2.2) Otherwise (i.e. $k_0 = 0$), the image $I_0$ will be distributed in cluster $V_m$ and $\phi(I_0, I_m) = k_0\theta$.

B. Extracting the Feature Region

For extracting the visual feature of image, the first step is standarized the image size (i.e. convert input image which has different size into the image have size $k \times k$), from that extracting the color feature of image. Because the image according to JPEG standard is described on color space YCbCr, so we need use YCbCr to extract specific information of the image. Setting Y, Cb, Cr as an intensity, Blue color, Red color respectively. In accordance [5], the convertibility from RGB to YCbCr as follow (with $R,G,B \in [0,1]$):

$$\begin{bmatrix}
Y \\
Cb \\
Cr
\end{bmatrix} = \begin{bmatrix}
65.481 & 128.553 & 24.996 \\
-37.797 & -74.203 & 112 \\
112 & -93.786 & -18.214
\end{bmatrix} \begin{bmatrix}
R \\
G \\
B
\end{bmatrix} + \begin{bmatrix}
16 \\
128 \\
128
\end{bmatrix}$$

According to [3], [4], the Gaussian transformation by human’s visual system is as follow:

$$L(x,y,\delta_p) = \frac{1}{10} [6G(x,y,\delta_p)+Y(\delta_p)+2G(x,y,\delta_p)*Cb+2G(x,y,\delta_p)*Cr]$$
with \( G(x,y,\delta_o) = \frac{1}{\sqrt{2\pi\delta_o^2}} \exp\left(-\frac{x^2+y^2}{2\delta_o^2}\right) \)

The intensity \( I_o(x,y) \) for color image is calculated according to equation:

\[
I_o(x,y,\delta_j,\delta_o) = \text{Det}(M(x,y,\delta_j,\delta_o)) - \alpha \text{Tr}^2(M(x,y,\delta_j,\delta_o))
\]

In there, \( \text{Det}(\bullet),\text{Tr}(\bullet) \) as Determinant and Trace of matrix, \( M(x,y,\delta_j,\delta_o) \) is a second moment matrix, be defined as follow:

\[
M(x,y,\delta_j,\delta_o) = \delta_o^2 G(\delta_o)^\text{T} \begin{bmatrix} L_x^2 & L_x L_y & L_y^2 \end{bmatrix} \begin{bmatrix} L_x & L_x & L_y \end{bmatrix}
\]

In there, \( \delta_j,\delta_o \) are the integration scale and differentiation scale, and \( L_{\alpha u} \) is the derivative computed the \( \alpha \) direction. The feature points of color image are extracted according to formula: \( I_o(x,y,\delta_j,\delta_o) > I_o(x',y',\delta_j,\delta_o) \), with \( x',y' \in A \)

\[
I_o(x,y,\delta_j,\delta_o) \geq \theta
\]

with \( A \) is the neighborhood of point \( (x,y) \) and \( \theta \) is a threshold value. A set of feature circles \( O_i = \{(o_i^1, o_i^2, ..., o_i^p)\} \) has a center is a feature points and a set of feature radius \( R_i = \{r_{i1}^1,r_{i2}^2,...,r_i^n\} \). Values of feature radius are extracted according to LoG method (Laplace-of-Gaussian) and have value in \([0, \min(M,N)/2]\), with \( M,N \) are the height and the width of image.

C. Creating a Binary Signature of the Image

Each feature region \( o_i \in O_i \) of the image \( I \) will be calculated histogram on the base of the standard color range \( C \), effectuating the clustering method relies on Euclidean measure in RGB color space to classify colors of every pixel on the image. Setting \( p \) is a pixel of image \( I \) and has a color vector in RGB as \( V_p = \{R_p, G_p, B_p\} \). Setting \( V_m = \{R_m, G_m, B_m\} \) is a color vector of a set of standard color range \( C \), so as to: \( V_m = \min\|V_p - V_m\|, V_p \in C \). At that time, at pixel \( p \) will be standardized in accordance with color vector \( V_m \). According to experiment, the paper will use the standard color range on MPEG7 to calculate histogram for color images on Corel database.

Setting \( o_i \in O_i \) \( (i = 1,...,N) \) is a feature circle of the image \( I \), histogram vector of the circle \( o_i \) according to MPEG7 standard will be as: \( H(o_i) = [H_x(o_i), H_y(o_i),...,H_z(o_i)] \), setting \( h_k(o_i) = H_x(o_i) / \sum_j H_x(o_j) \), standardize histogram vector is as \( h_k(o_i) = [h_k(o_i),h_k(o_i),...,h_k(o_i)] \). When, the binary signature to describe \( h_k(o_i) \) will be as \( B^k_i = b_i^1b_i^2...b_i^n \), with \( b_i^j = 1 \) if \( j = \left\lfloor (h_k(o_i') + 0.05) \times m \right\rfloor \), otherwise \( b_i^j = 0 \).

So, the signature to describe the feature region \( o_i \in O_i \) as

\[
\text{Sig}(o_i') = B^i_1B^i_2...B^i_n.
\]

For this reason, the binary signature of the image \( I \) will be as: \( S_I = \bigcup_{o_i'} \text{Sig}(o_i') \)

D. Creating S-kGraph

Input: Image database \( \mathcal{I} \) and threshold \( k\theta \)
Output: S-kGraph = \( (V_{SG},E_{SG}) \)

Algorithm 1. Create_S-kGraph(\( \mathcal{I}, k\theta \))

Begin
\( V_{SG} = \emptyset; \ E_{SG} = \emptyset; k_I = 1; n_I = 1; \)
For \( (\forall I \in \mathcal{I}) \) do
If \( (V_{SG} = \emptyset) \) then
Begin
\( I_o^n = I; \ r = k_I \theta; \)
Initialize cluster \( V_m = \langle I_o^n, r, \phi = 0 \rangle; \)
\( V_{SG} = V_{SG} \cup V_m; \)
End
Else
Begin
\( (\phi(I_o^n), k_I \theta) = \langle I_o^n, r, \phi = 0 \rangle; \)
If \( (k_I > 0) \) then
Begin
\( I_o^{n+1} = I; \ r = k_I \theta; \)
Initialize cluster \( V_m = \langle I_o^{n+1}, r, \phi = 0 \rangle; \)
\( V_{SG} = V_{SG} \cup V_m; \)
\( E_{SG} = E_{SG} \cup \{(V_{n+1},V_o) | \phi(I_o^n),k_I \theta \}; \)
\( n = n+1; \)
End
Else
Begin
\( \phi(I_o^n) = k_I \theta; \)
\( V_m = V_m \cup \langle I_o^n, k_I \theta, \phi(I_o^n) \rangle; \)
End
End
End

E. Image Retrieval Algorithm

Input: query image \( I_q \), S-kGraph = \( (V_{SG},E_{SG}) \), threshold \( k\theta \).
Output: set of similar image IMG.

Algorithm 2. Search-Image(\( I_q, S\)-kGraph, \( k\theta \))
Begin
\[\text{IMG} = \emptyset;\]
\[V = \emptyset; \phi_{\text{max}} = \phi(I_o, I_o') = \min\{\phi(I_o, I_o'), i = 1, \ldots, n\};\]
\[\text{For } (V_i \in V_{SG}) \text{ do}\]
\[\text{If } (\phi(I_o, I_o') \leq k\theta) \text{ then } V = V \cup V_i;\]
\[\text{For } (V_j \in V) \text{ do}\]
\[\text{IMG} = \text{IMG} \cup \{I'_i, I'_j \in V, k = 1, \ldots, |V_j|\};\]
\[\text{return } \text{IMG};\]

End.

IV. EXPERIMENTS

A. Model of Region-Based Image Retrieval

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{model.png}
\caption{The model of RBIR based on signature graph}
\end{figure}

Phase 1: Perform pre-processing

Step 1. Extract feature regions of the image in database into feature vector.
Step 2. Convert the feature vector of the image in the form of binary signature.
Step 3. Calculate the EMD distance of the feature signature of the image and insert into S-kGraph.

Phase 2: Implement Query

Step 1. For each query image, will extract the feature vector and convert into binary signature.
Step 2. Perform the process of binary signature retrieval on S-kGraph to find out the similar image.
Step 3. After we have the similar images, carry out an arrangement according to similarity from high to low and give the list of the image based on the similarity of binary signature.

B. The Experimental Results

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{result.png}
\caption{A sample result about image retrieval based on signature graph.}
\end{figure}

Figure 1. The model of RBIR based on signature graph

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{experiments.png}
\caption{a- Number of comparisons to create SG. b- The time to create SG}
\end{figure}

Figure 2. A sample result about image retrieval based on signature graph.

Figure 3. a- Recall. b- Precision

V. CONCLUSION

The paper gives the similarity evaluation method between two images on the base of binary signature and creates the S-kGraph to describe the relationship in of images. Since then, the paper creates the image retrieval system model on the base of feature regions, at that time to simulate the experiment on Corel’s image data classification. According to experimental result, the method of evaluation based on S-kGraph has speed up query similar images more than query in SSF (sequential signature file). However, the use of the feature of color will give an inaccurate result in the sense of image content. Therefore, the next development of the paper will extract the object on the image, since the paper will build the binary signature to describe objects as well as to describe the content of image.

REFERENCES

[1] Guang-Hai Liu, Jing-Yu Yang, “Content-based Image Retrieval Using Color Difference Histogram”, Pattern Recognition 46, 2013, pp.188-198
[2] Ilaria Bartolini, Paolo Ciaccia, Marco Patella, “Query Processing Issues in Region-Based Image Databases”, Springer-Verlag, Knowl. Inf. Syst 25, 2010, pp.389-420
[3] X. –Y. Wang, Jun-Feng Wu, Hong Ying Yang, “Robust Image Retrieval Based on Color Histogram of Local Feature Regions”, Springer Science + Business Media, Multimed Tools Appl 49, 2010, pp.323-345
[4] X.-Y. Wang et al., “Robust Color Image Retrieval Using Visual Interest Point Feature of Significant Bit-Planes”, Digital Signal Processing 23(4), 2013, pp.1136-1153
[5] Zhenjun Tang et al., “Robust Image Hash Function Using Local Color Features”, Int. J. Electron. Commun. (AEÜ) 67, 2013, pp.717-722.
[6] Corel Corp. http://www.corel.com.
[7] Y. Chen, Yibin Chen, “On the Signature Tree Construction and Analysis”, IEEE Trans. Knowl. Data Eng., vol. 18, no. 9, 2006, pp.1207-1224.
[8] Bahri Abdelkhalak, Hamid zouaki, “EMD Similarity Measure and Metric Access Method using EMD Lower Bound”, Inter. Journal of Computer Science & Emerging Technology, vol. 2, no. 6, 2011, 2011, pp. 323-33