Image Database Indexing and Retrieval Utilizing Global and Regional Based Feature Matching

Jehad ALNIHOUD

Al al-Bayt University, Mafraq, Jordan
Prince Hussein bin Abdullah Information Technology College
jehad@aabu.edu.jo

Abstract: A creative and efficient approach for CBIR (content-based image retrieval) is presented in this paper. Color and shape are considered and some of the primitive color features (statistical-based features) are calculated. After that and based on distance measure associated with SOM (self-organizing map), the image database was clustered and the BMU's (best matching units) were identified. Furthermore, the paper presents an enhanced edge detection algorithm to remove the unwanted pixels and to solidify objects within images. The proposed approach ease similarity comparison based on cross-correlation coefficient (ccc) which is used as new means of shape similarity between images. The proposed approach overcomes the computational complexity of applying bin-to-bin comparison as a multi dimensional feature vectors in the original color histogram approach and improves the accuracy of retrieval based on shape features.

Key words: CBIR, Color histogram, Cross correlation coefficient, SOM.

INTRODUCTION

Content-based image retrieval (CBIR) grabs the interest of many researchers and become one of the most active research fields nowadays. There are many popular CBIR systems include IBM’S QBIC project [FLI 95], VisualSeek [SMI 96], PicSOM [KOH 01], PicHunter[COX 00], and MIRROR [WON 05]. CBIR differs than many of other disciplines in computer vision because of the difficulty of its evaluation due to the fact that human subjectivity cannot totally ignored through the evaluation process [DEL 00] [LEW 00] [SAL 83]. New visual feature representations for image that provide an efficient discriminant for similarity queries has been the main interest for most of researchers in CBIR [GRA 95] [MIK 89][JAC 95][PAS 96][PEN 96][SMI 95]. The need for simple and efficient approaches to handle the color and shape retrieval problem in CBIR becomes one of the major challenges in this field of research.

The proposed approaches should consider the speed of performance, accuracy of retrieval, and the ability to achieve an accurate ranking for the retrieved images. In CBIR the color and shape-based retrieval are considered as the main component of the visual features which may consider in the retrieving process, while texture has no value if not associated with color. Color histogram (HC) is considered as one of the standard approaches for color-based retrieval. Moreover, there are many attempts to enhance this approach and to overcome some of the problems associated with the original color histogram approach. Color histogram approach relies on multi-dimensional feature vectors in which a bin-to-bin comparison is conducted. The computational complexity problem is obvious. Furthermore, incorporating color similarities into the distance function does not yield to a robust distance function that corresponds to the perceptual similarity of a color histogram.

In this paper, new approach for color-based image retrieval utilizing SOM with some primitive color features is presented. The proposed approach shows a significant improvement over the original color histogram approach. Moreover, as a second phase of the proposed approach a new and efficient algorithm to model and solidify objects is proposed and a new means of shape similarity based on cross-correlation is used.

The rest of the paper is organized as follows. Section 1 describes the proposed approach. Section 2 presents the performance evaluation. Section 3 illustrates experimental results. Section 4 draws conclusions.

1. Proposed approach

The proposed approach in this paper is a double
phases approach. In the first phase global color features are calculated and the image database is clustered using SOM. In the second phase objects within the images were modeled and cross correlation coefficient is used.

1.1. Phase I- Color-based retrieval approach

PicSOM deploy the SOM [KOH 01] as indexing technique to organize the feature vectors that due to the efficiency of SOM in organizing unsupervised statistical data. A SOM consists of a regular grid of map units. A model vector \( d_{v} \in \mathbb{R}^{d} \) is associated with each map unit \( v \). The map tries to represent all available observation \( d_{\mathbb{R}^{d}} \) with optimal accuracy.

The fitting of the model vectors is a sequential regression process, where \( t=0,1,2,\ldots, t_{\text{max}} -1 \) is the step index.

For each input sample \( x(t) \), first the index \( c(x) \) of the best matching unit (BMU) or winner model \( (t)c(x)m \) is defined by the condition:

\[
\forall v: \| x(t) - m_{c(x)}(t) \| \leq \| x(t) - m_{v}(t) \| \quad (1)
\]

After finding the BMU, a subset of the model vectors constituting a neighborhood centered around the BMU (node \( c(x) \)) are updated as:

\[
m_{v}(t+1) = m_{v}(t) + d(t;c(x),v)(x(t) - m_{v}(t)) \quad (2)
\]

where \( d(t;c(x),v) \) is the neighborhood function.

A decreasing function of the distance between \( v \)-th and \( c(x) \)-th units on the map grid. After the training phase, the BMUs partition the feature space into a set of Voronoi regions. The interior of each region consists of all points in the feature space that are closer to the respective BMU than to any other.

Four major global features are used in order to index the image database. Based on these features the (BMU) is identified with respect to the image query. This technique filters the image database and reduces the candidate images for the next stage. These features as follows:

1) Mean: The value of the Mean shows the general brightness of the image. As a general rule bright images has high mean, while dark image has low mean. The mean define as:

\[
M = \sum_{g=0}^{2^{l}-1} P(g) \quad (3)
\]

Where \( l \) is the gray level.

2) Standard Deviation: the standard deviation gives a clear idea about the image contrast. As a general rule high standard deviation means high image contrast, while small standard deviation means low image contrast. The standard deviation defines as:

\[
\sigma_{g} = \sqrt{\sum_{g=0}^{2^{l}-1} (g - \bar{g})^{2} P(g)} \quad (4)
\]

3) Energy: Energy shows how the gray level is distributed. The maximum value of energy is 1 and it become smaller as the pixel value distributed among the gray level. Energy defines as:

\[
\text{Energy} = \sum_{g=0}^{2^{l}-1} (P(g))^{2} \quad (5)
\]

4) Skew: Measures the asymmetry about the mean in the gray level distribution. Skew define as:

\[
\text{skew} = \frac{1}{\sigma_{g}^{3}} \sum_{g=0}^{2^{l}-1} (g - \bar{g})^{3} P(g) \quad (6)
\]

Feature vectors are merged and normalized. The normalized vectors of all images are fed into the map calculation algorithm which produces a map with hexagonal layout. Each cluster is represented by a feature vector pointing to its centre. Then the BMU of the query image is identified. The distance between the search image cluster and the neighboring clusters is identified through the calculation of the weight for all features based on the reciprocal value of the sum of distances.

For this purpose Euclidean distance function is used. At the end of this phase, the database images are filtered based on the cluster for which the search image and some of the neighboring clusters are belongs.

A metric space is a tuple \((M, F, I)\), where \( F \) is a feature set and \( M \) a metric on \( F \). Suppose that \( F_{j} \times F_{j} \) be the Cartesian product between features of the same space. Let \( M \) be a metric that calculates the similarity between a pair of given features, then:

\[
M_{j} : F_{j} \times F_{j} \rightarrow \mathbb{R}
\]

Such that:

1. \( M_{j}(x, y) \geq 0 \). Non-negativity
2. \( M_{j}(x, y) = 0 \), if and only if \( x = y \). Identity
3. \( M_{j}(x, y) = M_{j}(y, x) \). Symmetry
4. \( M_{j}(x, y) \leq M_{j}(x, y) + M_{j}(y, z) \). Triangle inequality.

The previous definition introduces an order relationship between images using a feature (1) and a metric \( M_{j} \). Better comparisons may be achieved using many features and a linear combination of different metrics. Let \( x, y \in I \) be images. Let \( E_{k} \)
be the feature extraction function of a feature \( k \) and \( M_k \) be a metric in the feature space \( F_k \). A similarity function for different features is defined as the linear combination of metrics \( M_k \) with importance factors \( W_k \):

\[
\text{Sim}(x, y) = \sum_k W_k M_k(E_k(x), E_k(y))
\]

(7)

### 1.2. Phase II- Shape-based retrieval approach

Many edge detectors are available to researchers. Marr and Hildreth convolve a mask over the image and label zero-crossings of the convolution output as edge points [MAR 80]. In [GRE 91], an approach relies on combination of contrast thresholding and an analysis of direction dispersion to find edges is presented. In [BAK 81], they label peaks in the magnitude of the first derivative of the intensity profile along a scan-line as feature points for matching. Other popular gradient edge detectors are the Canney, Roberts, Sobel and Prewitt operators [BAL 82]. Comparing objects based on edge detection does not yield to satisfactory results in most cases. Since any variation of image brightness affects the accuracy of comparing images based on edge detection. Moreover, the unwanted pixels in the image may affect the retrieval accuracy too. To overcome these problems an algorithm to filter the images at the pre-processing stage is proposed. Edges in images may be detected using one of the many available detectors. Based on empirical testing of, Prewitt operator is selected. In order to solidify the objects in the image query as well as the image database, a new algorithm presented in Figure 1. The proposed algorithm utilizes morphological operators in combination with edge detection operator and automatic cropping. The proposed algorithm enhances the retrieval accuracy dramatically due to its ability to:

- solidify objects within image.
- remove unwanted pixels within images.

**Object Modeling Algorithm(Image I)**

1. Apply Prewitt edge operator.
2. Use a proper threshold to remove unwanted edges.
3. Apply morphological (dilate) with a proper structuring element to fill in holes.
4. Automatic cropping to the image as follows:
   - scan image row by row
     - if \( I(x, y) = 0 \) \( \forall(x, y) \in r \) then crop the row
     - else continue
   - scan image column by column
     - if \( I(x, y) = 0 \) \( \forall(x, y) \in c \) then crop column
     - else continue

**Figure 1. Object modeling proposed algorithm (OMA)**

The following figure shows the system block diagram.

![System block diagram](image)

Since the objects in the image query as well as database were modeled and solidified based on the previous algorithm, new means of similarity based on cross correlation coefficient proves to be efficient and realistic. Cross-correlation is compares two different sequences. The cross-correlation function (CCF) of two sequences \( x[n] \) and \( y[n] \) is defined in terms of time averages by the following Equation.

\[
\Xi[m] = \lim_{n \to \infty} \frac{1}{2n+1} \sum_{i=-n}^{n} x[i]y[i+m]
\]

(8)

In signal processing, the cross-correlation (or sometimes "cross-covariance") is a measure of similarity of two signals, commonly used to find features in an unknown signal by comparing it to a known one. It is a function of the relative time between the signals, is sometimes called the sliding dot product, and has applications in pattern recognition and cryptanalysis.

The cross-correlation is similar in nature to the convolution of two functions. Whereas convolution involves reversing a signal, then shifting it and multiplying by another signal, correlation only involves shifting it and multiplying (no reversing). Sometimes it is preferable to express the cross correlation of two signals in terms of the cross-correlation coefficient. It was calculated by
normalizing the cross-correlation of the two signals with the power of the two signals i.e. by setting m = 0. The cross-correlation coefficient lies between -1 and +1, with zero indicating no correlation between the two signals.

$$\zeta_{xy}[m] = \frac{\mathfrak{X}_{xy}[m]}{\left| \mathfrak{X}_{xx}[0]\mathfrak{X}_{yy}[0] \right|^\frac{1}{2}}$$  (9)

2. Performance evaluation

The image comparison in the proposed CBIR module is based on query by example [PET 97]. The example image is analyzed and the necessary features are extracted in each phase then the compared with other database images. CBIR system like any other IR (information retrieval system) resolves queries in an approximate way, because the users are not specific about the precise results that should be delivered [YAT 04]. It is believed that what is important is the image retrieval module. Even so, it is good to evaluate the performance of that module. To measure the performance of any retrieval system, precision and recall are still the most prominent techniques to use. In [MIK 89] they present a framework to evaluate CBIR based on precision:

$$\text{Precision} = \frac{\text{No. of relevant retrieved images}}{\text{No. of all relevant images}}$$  (10)

For several queries average precision is preferable, which may define as:

$$\overline{P}(r) = \frac{\sum_{i=1}^{N_q} P_i(r)}{N_q}$$  (11)

where $\overline{P}(r)$ is the average precision at recall level r, $N_q$ is the number of queries, and $P_i(r)$ is the precision at recall level r for the i-th query.

3. Experimental results

3.1. Experimental results (color-based approach)

The proposed CBIR approach has tested on a general-purpose image database with 1000 images from COREL. These images represent 10 different categories with 100 images in each category. To evaluate the retrieval performance 10 images randomly selected from 5 categories with different semantics (Africa, Building, Dinosaur, Beach, and Vehicles). A retrieved image is considered as a correct match if and only if it belongs to the same category. As sample of retrieved results applying the proposed color-based approach, the top 9 returned images of 2 queries are shown in Figure 2. The query image is shown at the upper left corner.

3.2. Experimental results (shape-based approach)

The same image database from COREL is used to test the shape proposed approach. In order to test the reliability of the proposed approach, 20% of the tested images in each category are distorted or manipulated in a pre-processing stage (e.g. adding noise, rotate objects, translate objects, etc).

Table 1 lists the average precision for each category by applying the proposed method and color histogram method.

<table>
<thead>
<tr>
<th>Category</th>
<th>Proposed color-based approach</th>
<th>Traditional Color histogram</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beach</td>
<td>0.56</td>
<td>0.23</td>
</tr>
<tr>
<td>Africa</td>
<td>0.85</td>
<td>0.62</td>
</tr>
<tr>
<td>Building</td>
<td>0.84</td>
<td>0.35</td>
</tr>
<tr>
<td>Vehicles</td>
<td>0.78</td>
<td>0.19</td>
</tr>
<tr>
<td>Dinosaur</td>
<td>0.92</td>
<td>0.98</td>
</tr>
<tr>
<td>Average</td>
<td>0.79</td>
<td>0.474</td>
</tr>
</tbody>
</table>

Figure 3. Retrieval results of 2 queries  
Figure 4. Retrieval result of 1 query image (traditional color histogram approach)
Figure 5. Retrieval result of 1 query image (shape-based approach). The query image is shown at the upper left corner. 9 matches out of 9 retrieved images.

The overall precision is increased significantly as compared to the previous results. Table 2 shows the average precision when the double-phases approach (color and shape) is applied.

Table 2. Average Precision (Shape-based Approach)

<table>
<thead>
<tr>
<th>Category</th>
<th>Proposed color-shape approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beach</td>
<td>0.85</td>
</tr>
<tr>
<td>Africa</td>
<td>0.70</td>
</tr>
<tr>
<td>Building</td>
<td>0.90</td>
</tr>
<tr>
<td>Vehicles</td>
<td>0.95</td>
</tr>
<tr>
<td>Dinosaur</td>
<td>0.98</td>
</tr>
<tr>
<td>Average</td>
<td>0.88</td>
</tr>
</tbody>
</table>

4. Conclusion

A novel CBIR approach is presented in this paper. The approach shows a high level of retrieval accuracy. The approach utilizes database indexing based on SOM and filtered the image database based on some primitive features of the image. Furthermore, an enhanced edge detection technique is applied to detect objects in the image query as well as image database which make it easy to measure the similarity between the query image and the candidate images in the database.

The extensive testing of the proposed approach and the comparison with the color histogram approach proves that the proposed approach is able to improve the accuracy of retrieval dramatically. Furthermore, utilizing cross-correlation as new means of shape similarity to the solidified objects enhances the detection process and gives a high level of retrieval accuracy.

REFERENCES


