Interactive image contents search based on high dimensional information theory

CAO PENG FEI
Sanjiang University, Nanjing China, 210012, (e-mail: Pengfei.cao@outlook.com)
Corresponding author: Cao peng fei

ABSTRACT Content-Based Image Retrieval aims at developing techniques that support effective searching and browsing of large image repositories based on image features. The current Content-Based Image Retrieval suffer from the semantic gap, it often leads to distraction in the search. To overcome these disadvantages, Combine the theory of high dimensional information property, we propose different dimensional space’s vector angel cosine algorithm based on high dimensional information theory. The experimental results demonstrate the performance of the proposed algorithm.

INDEX TERMS Image Retrieval, Angel Cosine, High Dimensional Information, Feature Extraction

I. INTRODUCTION

The traditional content-based image retrieval method are all based on image’s bottom feature, that are image’s color feature, texture feature, spatial location feature and so on, then match the image by this features. Google and Baidu [1], [2] are all based on this theory and proposed the new image retrieval function. But as people are not satisfied with content-based image retrieval, and they more hope to retrieval according their understanding, which is semantic-based image retrieval. But the problem how to infer the image semantic from the image content is a difficult thing. Because it referred the content of image understanding. As the semantic gap between image content and semantic, it is hard to find a universal mathematical model, which mapped to image semantic from image feature. That is, switch from data to knowledge is a difficult thing. Many techniques have been developed for text-based information retrieval and they proved to be highly successful for indexing and querying web sites [3]–[5]. Although it’s difficult, Not only Cao wenming [6] and He tiancheng have propose high dimensional geometry theory to applicable for speech emotion recognition, but also use geometric algebra do image segmentation [7], relative feedback [8], and build complex no-linear model [9]. As image segmentation is a open problem, so to the general purpose’s image retrieval [10], it is difficult to find a image segmentation arithmetic which is commonly used and fit for people’s understanding, what’s more, the image segmentation has a large computation and is difficult to realize real-time computation. But building complex non-linear model is also have difficult, because after extracted the bottom feature, the data amount decreased dramatically, so build the semantic model is very hard in the real word. Consequently, the current means it is a good choice to utilize Human-computer interaction. A new pattern recognition method—biomimetic pattern recognition theory, in recent years [11]–[15] is already have a good application in face detection, face recognition, speech recognition, and so on. It proposes a new image retrieval arithmetic which based on image content. This paper combines relative high dimensional information theory, proposes the new image search method which is based on image content. This algorithm is based on one-sided aspect that people are interested in. That is, how to intercept the area people interested in from a complex image. This paper adopts the area’s relative high dimensional information character, processing the image retrieval. Through the experiment, we concluded that the method effectively improve the accuracy of image retrieval and decrease the algorithm time.

II. METHOD

A. THE RELATIVE THEORY OF HIGH DIMENSIONAL GEOMETRY

The theory of high dimensional biomimetic information geometry and compute process are all start from the concept of high dimensional geometry figures. Apply the point in space to describe subspace, by means of plane geometry simple operation combination iteration realize the complex operation of high dimensional space, this chapter use Linear algebra tool to briefly describe space geometry system, and come up with the classification theory of high dimensional information geometry feature.

Definition 1 [16]: suppose \( (V, \langle \cdot, \cdot \rangle) \) is an n-dimensional Eu-
clidean vector space. Affine space which takes V as adjoin vector space is called n-dimensional Euclidean space. Distance between any two points \( P, Q \) defined as bellow:

\[
(P, Q) = \sqrt{(P\bar{Q}, \bar{P}Q)}
\]  

(1)

\( \mathbb{R}^n \) is an n-dimensional Euclidean vector space, meanwhile, \[ x = (x_1, \cdots, x_n) \in \mathbb{R}^n, \ y = (y_1, \cdots, y_n) \in \mathbb{R}^n, \] the distance between the two points defined as bellow:

\[
|x - y| = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}
\]  

(2)

In \( (\mathbb{R}^n, \langle, \rangle) \), the distance between \( x, 0 \) is: \[ |x| = \langle x, x \rangle^{\frac{1}{2}}, \] which also called the size of the vector \( x \). When \( x \neq 0 \) and \( x \neq 0 \), the included angle cosine of the two vectors \( x, y \) is:

\[
\cos \theta = \frac{\langle x, y \rangle}{|x||y|}
\]  

(3)

Where \( |x - y| \) is the length of line segments \( xy \) or vector \( x - y \). Now, \( \mathbb{R}^n \) can be called n-dimensional Euclidean space.

1) The Basic Theory Of Using Different Dimensional Space To Define The Angels

Let \( V \) be an r-dimensional subspace, \( U \) be an s-dimensional subspace in \( \mathbb{R}^n \).

**Step 1:** Let \( S^{n-1}_1(0) \) be a unit sphere in \( \mathbb{R}^n \), \( f(x, y) = \langle x, y \rangle, x \in V \cap S_1^{n-1}(0) \) and \( y \in U \cap S_1^{n-1}(0) \) are continuous in compact set \( (V \cap S_1^{n-1}(0)) \times (U \cap S_1^{n-1}(0)) \), thus, where \( x_0 \in V, y_0 \in U \), obtain the maximum value, namely, for any \( x \in V \cap S_1^{n-1}(0) \) and \( y \in U \cap S_1^{n-1}(0) \), then \( \langle x_0, y_0 \rangle \geq \langle x, y \rangle \). Let \( \theta \) be the minimum acute angle between \( x_0 \) and \( y_0 \), then \( \cos \theta = \langle x_0, y_0 \rangle \). Thus, \( \theta \) can be defined uniquely.

As we can concluded that the orthogonal projection from \( X_0 \) to \( U \) is linear relationship with \( y_0 \). Conversely, it is also true.

\[
\langle x_0 \rangle \perp U \Leftrightarrow \langle y_0 \rangle \perp U \Leftrightarrow \langle x_0 \rangle \perp \langle y_0 \rangle \perp V
\]  

(4)

The uniqueness about \( x_0 \) and \( y_0 \). We take two vectors \( v \in V, |v| = 1, v \perp x_0 \) and \( u \in U, |u| = 1, u \perp y_0 \), arbitrary , then(Fig. 1),

\[
\langle \langle x_0, v \rangle \rangle \cap \langle \langle y_0 \rangle \rangle \perp U = \langle \langle v \rangle \rangle \perp U \cap \langle \langle y_0 \rangle \rangle = \langle \langle y_0 \rangle \rangle \perp V
\]  

(5)

The uniqueness about \( x_0 \) and \( y_0 \). We take two vectors \( v \in V, |v| = 1, v \perp x_0 \) and \( u \in U, |u| = 1, u \perp y_0 \), arbitrary.

Let \( \beta \) be the included angle between \( x \) and \( y \), we can define the cosine function of \( \theta \) as:

\[
\cos \theta = \frac{\langle x, y \rangle}{|x||y|}
\]  

(6)

\[
\langle x, y \rangle = \alpha_1 x_0 + \alpha_2 v, \alpha_1^2 + \alpha_2^2 = 1
\]  

(7)

Let \( y = \beta_1 y_0 + \beta_2 u \), making \( x - y = \alpha_1 x_0 + \alpha_2 v - \beta_1 y_0 - \beta_2 u \) vertical to \( \langle \langle y_0, u \rangle \rangle \). Then,

\[
\langle x - y, y_0 \rangle = \alpha_1 \langle x_0, y_0 \rangle - \beta_1 \langle y_0, y_0 \rangle = \alpha_1 \cos \theta - \beta_1 = 0
\]

\[
\langle x - y, u \rangle = \alpha_2 \langle v, u \rangle - \beta_2 \langle u, u \rangle = \alpha_2 \langle v, u \rangle - \beta_2 = 0
\]

\[
\Rightarrow y = \alpha_1 \cos \theta y_0 + \alpha_2 \langle v, u \rangle \]

FIGURE 1: two vectors accute angle geometry

Furthermore, the formula with which calculating cosine value of angle between X and \( \langle y_0, u \rangle \) is shown below:

\[
\langle \langle x, y \rangle \rangle \perp U \cap \langle \langle y_0 \rangle \rangle = \langle \langle y_0 \rangle \rangle \perp V
\]  

(8)

If the result of the equation is \( \cos \theta \), the necessary and sufficient condition is \( \alpha_1^2 \cos^2 \theta - \alpha_2^2 \langle v, u \rangle^2 = 0 \), that is \( \alpha_2 = 0 \) or \( \langle v, u \rangle = \cos \theta \), \( \alpha_2 = 0 \) shows that \( x_0 \in V, y_0 \in U \) is the unique vector that can meet:

\[
\cos \theta = \langle x_0, y_0 \rangle > \langle u, v \rangle
\]  

The \( \langle v, u \rangle = \cos \theta \) show that the angle between any non-zero vector in \( \langle x_0, v \rangle \) and plane \( \langle y_0, u \rangle \) is identical.

**Step 2: Inductive Definition**

If for every \( x \in V \cap S_1^{n-1}(0), y \in U \cap S_1^{n-1}(0) \), there exists an element \( x_1 \in V \cap S_1^{n-1}(0), y_1 \in U \cap S_1^{n-1}(0) \) such that:

\[
\cos \theta_1 = \langle x_1, y_1 \rangle \geq \langle x, y \rangle
\]  

(9)

The definition of \( \theta_1 \) is unique, and it is independent of \( x_1, y_1 \).

For every \( x \in V \cap S_1^{n-1}(0), y \in U \cap S_1^{n-1}(0) \), there is a function \( f(x, y) = \langle x, y \rangle \), which can get maxima at the point of \( x_2 \in V \cap S_1^{n-1}(0) \cap \langle x_1 \rangle, y_2 \in U \cap S_1^{n-1}(0) \cap \langle y_1 \rangle, r > 1, s > 1 \). We can define the cosine function of \( \theta_2 \) as:

\[
\cos \theta_2 = \langle x_2, y_2 \rangle
\]  

(10)

Then the equation \( \theta_1 \leq \theta_2 \). The definition of \( \theta_1 \leq \theta_2 \) that \( \theta_1 = \theta_2 \) is unique, and it is independent of \( x_2, y_2 \).

Let be \( k = \min(r, s) \), for any fixed \( l \leq k, \) \( \theta_1 \geq \cdots \geq \theta_2 \geq \theta_1 \) can be defined unique, which makes \( \theta_1 > \theta_1 - \theta_2 \) or \( \theta_1 = \theta_1 - \theta_2 \). \( 2 \leq i \leq l - 1 \). Let \( 2 \leq i \leq l - 1, y_i \in U \cap S_1^{n-1}(0), \) which makes

\[
\cos \theta_i = \langle x_i, y_i \rangle, 1 \leq i \leq l - 1
\]  

(11)

As the function \( f(x, y) = \langle x, y \rangle \) can get the maxima on condition of \( x \in V \cap S_1^{n-1}(0) \cap \langle x_1, \cdots, x_{i-1} \rangle \) and \( y \in U \cap S_1^{n-1}(0) \cap \langle y_1, \cdots, y_{i-1} \rangle \). That is there \( x_i, y_i \) which makes the \( \langle x_i, y_i \rangle \) get the maxima.

Take a notice to \( \theta_i > \theta_{i-1} \) or \( \theta_i = \theta_{i-1} \), whether \( x_i, y_i \)
exist only or not. This process stop at \( l = k \). Therefore,
\[
1 \geq \cos \theta_1 \geq \cos \theta_2 \geq \cdots \geq \cos \theta_k \\
0 \leq \theta_1 \leq \theta_2 \leq \cdots \leq \theta_k \leq \frac{\pi}{2}
\]
This angle of \( \theta_1, \theta_2, \ldots, \theta_k \) is defined by the subspace of \( V \) and the subspace of \( U \) which is called the angle of subspace.

**Step 3: The continuous and equal angle.**

Let \( p, 0 \leq p \leq k - l \) is an integer, which makes \( \theta_{l+p} < \theta_{l+1} \) \( \leq \cdots \leq \theta_{l+p} \leq \theta_1 \) (if \( l + p = k \), it will stop at \( \theta_{l+p} \)).

In this case, the subspace could be denoted by
\[
V_i = \langle \{x_i, x_{i+1}, \ldots, x_{i+l}\} \rangle \\
U_i = \langle \{y_i, y_{i+1}, \ldots, y_{i+l}\} \rangle
\]

Where
\[
\langle \{x_i, x_j\} \rangle = \langle \{y_i, y_j\} \rangle = \delta_{ij}, l \leq i, j \leq l + p
\]
\[
x_i \perp y_j, l \leq i \leq l + p - 1, l + 1 \leq l \leq l + p
\]
\[
y_i \perp x_j, l \leq i \leq l + p - 1, l + 1 \leq j \leq l + p
\]

For any vector \( x \in V_i \cap S_1^{n-1}(0) \), it can be expressed as
\[
x = \alpha_1 x_1 + \alpha_{1+1} x_{l+1} + \cdots + \alpha_{l+p} x_{l+p}
\]

The orthogonal projection between \( x \) and \( U_i \) is
\[
y = \frac{x}{\|x\|} = \left(\cos \theta_i \left(\alpha_1 y_1 + \alpha_{1+1} y_{l+1} + \cdots + \alpha_{l+p} y_{l+p}\right)\right)
\]

The angle between \( x \) and \( y \) is \( \theta_i \), which can be expressed as
\[
\left\langle \begin{array}{c} x \\ y \end{array} \right\rangle = |y| = \cos \theta_i \left(\alpha_1^2 + \alpha_{1+1}^2 + \cdots + \alpha_{l+p}^2\right)^{\frac{1}{2}} = \cos \theta_i
\]

Contrarily, \( \frac{x}{|x|} \in U_i \cap S_1^{n-1}(0) \), and the orthogonal projection between it and \( V_i \) is \( \frac{x}{|x|} \).

In the version, the image \( A \) could be expressed as vector \( A \). The angle between is \( \langle A, A \rangle = \frac{A \cdot A}{|A||A|} = 1 \). The formula with which calculating cosine value of angle between two vectors is \( \cos(A, B) = \frac{A \cdot B}{|A||B|} = 1 \). When \( \cos(A, B) = \frac{A \cdot B}{|A||B|} = 1 \), the vector \( A \) is equal to vector \( B \). That is the image \( A \) and image \( B \) has the minimum difference.

The algorithm in this paper is base on the cosine value of the image. We define a parameters \( k \). When the cosine value is \( \cos(A, B) \geq k \), it means that the two images is similar. Contrarily, if \( \cos(A, B) < k \), the two images are different.

**B. THE INTERACTIVE IMAGE RETRIEVAL METHOD**

This algorithm starts from people’s general law recognition, apply the law of things to image retrieval. Steps of the algorithm is as follow:

**Step 1:** Take an image \( A_{MN} \) intercepting the region of interest in the image, and marking it as \( A_{mn} \)

\[
A_{mn} = \begin{bmatrix}
a_{11} & a_{12} & \cdots & a_{1n} \\
a_{21} & a_{22} & \cdots & a_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
a_{m1} & a_{m2} & \cdots & a_{mn}
\end{bmatrix}
\]

To process the other pictures in the image library, we take a picture \( B_{MN} \) from the image library random, then divide it into several blocks[9]. First, the image is divided into 4 blocks (2 rows x 2 columns). The result is shown in Fig. 1. If we use this method, a complete image may be divided into several meaningless image blocks. This result is unfavorable for image retrieval. In this paper, we propose overlapping block method. On some occasions, this method can avoid the above problem. The principles of image block are as follows: We can use the first step of \( A_{MN} \) as a template. The overlapping block method can be applied to \( B_{MN} \). We suppose the total number of blocks is \( p(q \) may be greater than or equal to \( \frac{M \times N}{m \times n} \)).

**Step 2:** We can apply the theory of High Dimensional Space Geometry for each image sub-block. The geometric mapping can be applied to the image sub-block. That is feature extraction [7]. Each sub-block after feature extracting is marked as \( D_{l}^{x \times e} \), \( l = 1, 2, \ldots, q \). At this time, \( w \times e \) should be less than \( m \times n \). The query image \( A_{mn} \) after feature extracting is marked as \( A_{x \times e} \).

**Step 3:** We can reshape the matrix \( A_{x \times e} \) according to column into a vector \( a_{x \times e} \), the dimension is \( W \times e \). Each sub-block of the image \( B \) which is come from the image library can be marked as \( D_{l}^{x \times e} \), \( l = 1, 2, \ldots, q \). By the same method, the \( D_{l}^{x \times e} \) can be reshape to \( d_{l}^{x \times e} \), \( l = 1, 2, \ldots, q \).

**Step 4:** Similarity Measurement. We intercept a main region which can instead of the query image. This can reduce the computing time. To measure the similarity between a database image and the query image, we used included angle cosine, which is given by
\[
T = \max \left( \lambda \left( a_{x \times e}, d_{l}^{x \times e} \right) \right)
\]
where \( \lambda(a, d) \) is included angle cosine of two vectors, \( q \) is the number of image sub-block.

We can take an image shareholding \( \delta \). If \( T \geq \delta \), then the database image is similar to the query image. In other words, there is a region in the database image that the user is interested in it. If \( T < S \), here is no region in the two images.

**III. EXPERIMENT AND RESULTS**

**A. EXPERIMENTAL PLATFORM**

- Hardware environment: CPU (R) T2400, dominant frequency 1.83G, internal storage 1.00GB.
- Software environment: windows XP operating system, image form platform matlab7.5(R2007b).

The data this experiment used are ORL database which are used for face recognition, this database include 40 different people image, and every people’s gesture is different, we use this database for the first retrieval. This paper is also applying the algorithm to COREL image database. This database is more comprehensive which contains figure, animals, scenery, vehicle, food and so on. This images are universally used in
the worldwide. Bellows are the retrieval results.

### B. ORL IMAGE [17] HAVE THE SIMILAR RETRIEVAL RESULT.

First, as the requirement of web retrieval, many users need to retrieval out one person. That is, when he input his one image, all the images about him can be retrieved out. Take this into consideration, this paper apply the proposed retrieval method to ORL face library, the experiment results indicate that the proposed method has an advantage on recall and precision ratio, the query results is shown in Fig. 2. From the query results, we conclude that the proposed method have the ability to retrieval all the images about the person.

In the retrieval to the face library, there is no more segmentation on the image. Because the face is intact, and the resegmentation may result to blockette hard to recognize, such as eyes, nose etc. In this case, the matching result is similar with some organ, rather than the whole face. So, when segmentation blocks can representation one thing, there is no need to resegmentation. And resegmentation has no meaning.

Based on the consideration, this paper utilizes the overlapping partition method to split an image. Only the segmented images is undamaged, can realize the requirement.

Above is just by ORL face library, to verify the algorithm, the changes of this images is small, so may haven’t much persuasion.

### C. COREL IMAGE [18] HAVE THE SIMILAR RETRIEVAL RESULT.

The image library used by the experiment includes 23 images. Among them, there are 6 people face images, 5 pillar images, 4 dinosaur images and the other flower and scenery images. They are selected from Corel image library. The search time is 33 seconds. They are arranged by similarity, and the 9 images which have higher similarity show in right of user interface. Arranging the images by similarity from high to low. Fig 3 shows 8 images from the image library used for this experiment. The image opened by the user displayed in the upper left corner of the user interface. The user can draw a black rectangle on the image with the mouse. The image in the rectangle is what the user wants to search.

Nine images with the top similarity are displayed by computer in the right side of the user interface, and they are displayed by similarity from left to right, high to low.

The images searched in the experiment include 1) top of the pillar; 2) bottom of the pillar; 3) blue sky; 4) flower; 5) ordinary human face; 6) special human face;

1) Top of the pillar

The experimental result can be seen in Fig. 4. We want to search the pillar, hoping to find some similar parts to the pillar. 10,11,22,4.jpg have the similar pillars in the image library. These images are the goals we hope to find out. We can see that the 4 images have the top similarities. The experimental result correspond to expectation.

2) Bottom of the pillar

The experimental result is shown in Fig. 5. This time, we searched the bottom of the pillar. The result is quite different from the first one. In the 9 images which has the top similarity, 6 of them contain human faces or flowers.

In my opinion, it is because the bottom of the pillar looks a little like human face and flower.

What we can see from the experimental result is that although the user opens the same image, but if the user choose to search different contents, the results may be quite different.
3) Blue sky
Fig. 6 shows the experimental result. The blue sky has the following features: the pixel values are not all the same, but the distinction are very small.

There are 7 images which contained similar blue sky in the image library. We can see that these images have the highest similarity. The result is as desired.

4) Flower
The experimental result is shown in 7. In the 9 images which have the highest similarity, the first 3 images contain human faces, and other 3 images contain flowers. What we searched is part of the flowers, but there are similarity between them, so the search result is reasonable.

5) Ordinary face
The experimental result is shown as Fig. 8. There are 6 images which contains face in the image library. From the experimental result we can see that 5 of them have high similarity.

We also can see that the flower image is ranked the fourth. It can be interpreted that although the flowers and the face are not the same thing, but there are similarity in appearance, so that they may also have high similarity.

6) Special face
The experimental result is shown as Fig. 8. The human we searched has been painted some colors which is different from complexion. It is a very special face.

We can see that all the 6 images which contains face are displayed in the right side of the user interface.

The experimental result shows that although the face we choose is not an ordinary one, but by the algorithm we still can search out some human images.

IV. CONCLUSIONS
This paper is based on the high dimension biomimetic information geometry theory. Combining with the general laws of things people recognized, proposed the new image segmentation method and new similarity measurement method. The experimental result shows that, the method proposed can meet the requirement that people hope to the diversity of the image retrieval. What’s more, this paper utilize interactive fork value method, make the accuracy of the image similarity more higher.
V. ACKNOWLEDGEMENT

This work was supported by the Natural Science Foundation of the Jiangsu Higher Education Institutions of China under Grant NO.17KJB520032.

REFERENCES