

An Approach of Image Retrieval for Natural Scenes based on Local Features and Topology

Jun YANG¹

Abstract. Image retrieval problem is a hotspot research filed in recent decades since it has a wide range of applications. Topology studies which used for image retrieval has rare research. This paper has characterized the departure from the image, fully tap the characteristics of the image itself, using image feature characteristics of local feature points to achieve the image content with regional dimensions. Marking the completion of the image content, the use of two images may be compared with similar characteristic sequence identification, signature sequence with the greatest similarity, which is considered to be similar. The approach which including topology theory can overcome the natural image retrieval at the local characteristics of defects, which can protect the global sequences, has a good characterization. Methods for the analysis of natural images operation, experimental results show that the combination of local feature topology method can overcome the mere use of local characteristics identified defects, to achieve the purpose of image retrieval.

Keywords: image retrieval; topology; local features.

1 Introduction

According to the survey, nearly 80 percent information obtaining by humans are from the vision, it is a very critical way of human beings in knowing the objective world [1]. In the information age, image information as one of the most common multimedia information which provide a more intuitive, vivid, rich description has play more important role in acquiring knowledge. Digital cameras, scanners and other image equipments make digital images at an alarming rate to be created, while the large-capacity storage technology, database systems and Internet tech-

¹ J. Yang (✉)

Shanghai University of Electric Power, 200090, Shanghai., China
e-mail: yangjun@shiep.edu.cn

nology development, but also for its spread provides a powerful medium. So increasing the number of images, image information organization of the disorder is increasingly prominent, and therefore establishes an accurate and efficient information management system has become an image problem to be solved, this demand is to promote the image retrieval technology [2].

As an excellent local feature descriptors, SIFT operator is widely used in object recognition, image registration, indoor scenes positioning occasions [3]. Sivic and Zisserman SIFT features will also be applied to content-based video retrieval system Video Google, as SIFT in image retrieval application and development of the field has opened up a broader view [4].

Image matching technology has become computer vision and digital image processing and other areas of research focus, at present, the image matching methods are mainly based on pixel and two kinds of feature-based approach. Lowe proposed a new point feature extraction algorithm-SIFT algorithm group solved the scenes partial occlusion, rotation scaling, image distortion caused by changes in viewpoint and other issues [5]. Literature [6] on the SIFT algorithm has been improved through the feature point neighborhood gradient statistical analysis for image matching. Literature [7] By DoM (mean difference) scale space using the integral image and integral histogram detection feature points to generate a feature descriptor. Literature [8] proposed a SIFT feature based on gray information descriptor.

Therefore, SIFT feature descriptor SIFT algorithm is the core algorithm processing time and accounted for most of the performance of the algorithm and real-time impact is huge. This paper presents a rotation invariant texture of the SIFT descriptors for the generated key feature points, the key point is calculated for each image region around the rotation invariant texture features, will be integrated into the RIT characterization SIFT descriptor vector, more comprehensive and describe the image information. Meanwhile, the full rotation invariant texture features is used to further reduce the time sequence image and the reference image matching computational complexity. Textured established local feature descriptor, containing the positioning error for feature matching provides better fault tolerance to a certain extent, reducing the false match rate and improving the matching results.

This paper is organized as follows. Feature descriptor given in the next section and its improvement strategy, the feature descriptor matching is described in section 3, in the last section, experiment and its results are given.

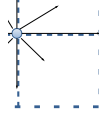
2 Feature Descriptor

For the content of natural images, the background is relatively complex, its computational time complexity is higher, is required for the characterization of feature

descriptor can improve execution efficiency. The following descriptions have two aspects.

2.1 Simplified characterization

In method [1], a scale-normalized image region is represented with the concatenation of gradient orientation histograms relative to several rectangular surrogates.



(a) 8 components

(b) 4 components

Fig.1. Descriptor of a sub-region

SIFT descriptors have 8 components and a key point descriptor includes a 128-dimension vector. In this paper, we use 4 components of sub descriptor for a sub-region is considered based on two points, one is that the performance can be improved by using properties of tag in estate image, the other is that we can decrease components size for speed produce.

2.2 Region of Interest

Tag regions are characterized by a high density of edges. Neumann pointed out the gradient in text regions exhibit an anti-parallel property. Based on these observations, we define text-like components using connected component (CC) based methods [9].

$$CC_{density}^{\theta} = \frac{\sum_{(x,y)} G(x,y)_{\theta}}{|CC|} \quad (2.1)$$

Where θ is $[i/4 * \pi]$, $i = 0, 1, 2, \dots, 8$, $G(x, y)$ denotes the gradient magnitude obtained from the Gaussian derivative of the gray scale image. In order to classify the CCs into text and non-text classes, Sort $CC_{density}^{\theta}$ into sequence by different channel θ . It should be symmetry and has max value in direction $\theta = 0, 1/2\pi$.

Owing to its simplicity, we choose the block-based Otsu threshold technique to binarize the gray-scale image patch described by each CC where the image patch is sub divided into 4×4 blocks.

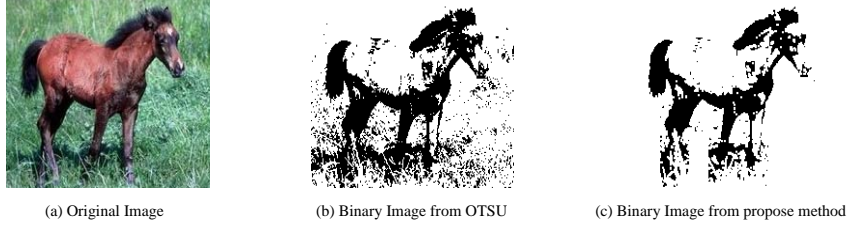


Fig2. Example image and its binary image

3 Feature descriptor matching

Obtained through the SIFT features described on the basis of experimental observations is shown in Figure 3.

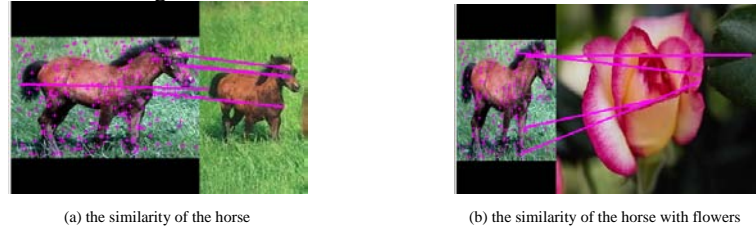


Fig3. Similar characteristics with corresponding connection

In Figure 3, the typical comparison of the similarity, the similarity can be seen that the local characteristics of the horse in the order of the sequence has been, and the flowers appear in the order sequence of the position hint reversed, thus the local characteristics of the same type that has a certain sequence order, retrieve images and between images to be retrieved with the appropriate characterization of sequence numbers can be completed through the description of a sequence. Special characteristic sequence coordinates are defined as follows:

$$F_1 = \{(x, y) \mid x \in [1..M_1], y \in [1..N_1]\} \quad (2)$$

$$F_2 = \{(x, y) \mid x \in [1..M_2], y \in [1..N_2]\} \quad (3)$$

In which, M_1 , N_1 is the length and width of the original image respectively. M_2 , N_2 is the length and width of the retrieval image.

Definition mapping sequence is as follows:

$$M = \{(x, y) \rightarrow (x^o, y^o) \mid (x, y) \in F_1, (x^o, y^o) \in F_2\} \quad (4)$$

The M sequence is defined to achieve the opposite.

When $x_i \leq x_{i+1}$ and $x_i^o \leq x_{i+1}^o$, the same order sequence result is 0, otherwise is 1. Symbolized as

$$s = \begin{cases} 1, & \text{if } x_i \leq x_{i+1} \ \& \ x_i^o \leq x_{i+1}^o, \\ 0, & \text{else} \end{cases} \quad s \in S \quad (5)$$

S is the last sequence result.

$S = \{S^0, \dots, S^0 \dots, S^1 \dots\}$, in which $S^0 = \{0 \dots 0 \dots\}$, $S^1 = \{1, 1 \dots\}$. When count of S^1 bigger than S^0 , $size(\sum S^1) / size(\sum S^0) > T$, it is considered as the candidate retrieval result.

For features of copyright tag, they are always distributed in horizontal and vertical direction. In order to speed compute rate, Euclidean distance was applied for features. It uses corner and edge detector to filter features. After filtered, the features with short energy corner response were removed. By using this method, time of matching the feature is shorted. Euclidean distance is used for comparing two vectors. But in our method, we use Hellinger distance for measure similar.

It is well known for classifier application such as text and image classification especially for unbalance data set, Hellinger distance measure is more validity of the established as the choice of suitable features.

The Hellinger distance is another way to reflect the distance of two distributions. Suppose in metric space (Θ, \mathcal{A}) , P, Q are two vectors. For Hellinger distance between them is defined as:

$$H(P, Q) = \sqrt{\sum_{\phi \in \Theta} (\sqrt{P(\phi)} - \sqrt{Q(\phi)})^2} \quad (6)$$

where

$$\begin{cases} 0 \leq H(P, Q) \leq \sqrt{2} \\ H(P, Q) = H(Q, P) \end{cases} \quad (7)$$

Between the two vectors, $H(P, Q)$ is non-negative and symmetric value. In this paper, we use Hellinger distance instead of Euclidean distance.

4 Experiments and Results

4.1 Natural image set

In this study, experimental use of animals and plants to experiment, image size variable, as described below.

Tab.1. Test Image Set Description

Image category	The amount of Searching	The amount of retrieval results	Total
horse	40	360	400
flower	40	460	500
elephant	40	360	400
total	1300		

4.2 Retrieval Processing

First on the characteristics of the image to be retrieved to obtain, using image features retrieved when the comparison result meets certain threshold is considered the same. Feature use four directions SIFT feature, set the two compare features directly adjacent threshold range is 0.618, the average test results for cross-validation.

Test procedure is as follows:(1) select retrieval test images;(2) the search image gradation is processed;(3) obtaining four directions features of SIFT;(4) images based on SIFT access is be retrieved;(5) the sequence value after comparing SIFT is obtained;(6) a sequence value, determined using the matching algorithm, the number ratio of 0.618.

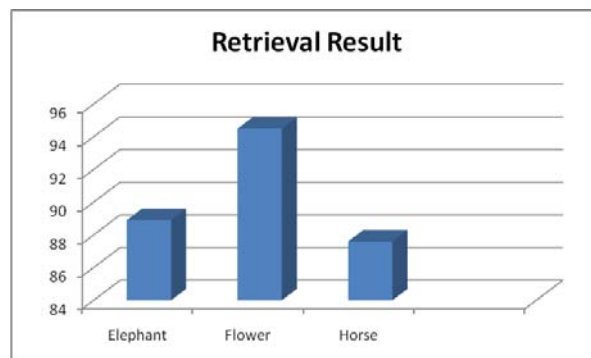


Fig.4. The probability distribution of similar search

Figure 4 shows the test results of different images, the retrieval rate is higher than spend horses and elephants, the main reason lies in the flower structure is relatively simple and relatively gather local features. Experiment, the characteristics of the SIFT time evaluation, roughly the test time shown in Table 2.

Table.2. Time Efficiency Comparison

Local features	8 Direction	4 Direction	Efficiency Improved
Average time	220ms	184ms	17%

4.3 Results

In the present study, we characterized the partial images are analyzed, including homogeneous and heterogeneous image by comparing the sequence to achieve the overall definition of the image information described in order to overcome the problem of local feature can not be compared. Contributions of this paper are as follows:

- (1) Image matching techniques have been proposed, the overall characterization of the way through the sequence of introduction.
- (2) It can not be effectively overcome local feature similar to that described image defects, making local features and global features combined; improved SIFT feature access methods, making efficiency can be improved, suitable for natural image retrieval services.

Experimental results show that the proposed method used in the natural image retrieval is effective, especially for local features can not describe the case of similar images, more about the global topology of the contents of the image to be improved.

Acknowledgement

This work is supported by National Nature Science Foundation of China (No. 61073189, 61202369). The authors are grateful for the anonymous reviewers who made constructive comments.

References

- [1] Koenderink J J.(1984) The structure of images[J]. Biological cybernetics, 50(5): 363–370.
- [2] Qianni Zhang and Ebroul Izquierdo.(2013) Histology Image Retrieval in Optimized Multifeature Spaces. IEEE journal of biomedical and health informatics, 17(1):240–249.
- [3] Jian Wu, Zhiming Cui(2013). A Comparative Study of SIFT and its Variants. Measurement Science Review. 13(3):122–131.
- [4] Joan Bruna, Stephane Mallat(2013). Invariant Scattering Convolution Networks, IEEE transactions on pattern analysis and machine intelligence , 35(8):1872-1886.
- [5] Lowe, D.G(2004). Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision, 2(60):91-110.
- [6] Liu Z, Xing B, Chen Y(2012). An efficient parallel PCA-SIFT algorithm for multi-core processor. Journal of National University of Defense Technology, 34(4): 103-107.
- [7] Wang Y W, Yu H L(2013). Image Registration Method Based on PCA-SIFT Feature Detection[J]. Advanced Materials Research, 712: 2395-2398.
- [8] Pang Y, Li W, Yuan Y, et al(2012). Fully affine invariant SURF for image matching[J]. Neurocomputing, 85: 6-10.
- [9] He L, Chao Y, Suzuki K, et al(2001). Fast connected-component labeling[J]. Pattern Recognition, 42(9): 1977-1987.