Image Retrieval Method Based on IPDSH and SRIP

Xu Zhang, Baolong Guo, Yunyi Yan, Wei Sun and Meng Yi
Institute of Intelligent Control and Image Engineering, Xidian University
Xi'an, Shaanxi 710071, China
[e-mail: xuzhang@mail.xidian.edu.cn, {blguo, yyyan, wsun }@xidian.edu.cn, yimeng0120@gmail.com]
*Corresponding author: Baolong Guo

Abstract

At present, the Content-Based Image Retrieval (CBIR) system has become a hot research topic in the computer vision field. In the CBIR system, the accurate extractions of low-level features can reduce the gaps between high-level semantics and improve retrieval precision. This paper puts forward a new retrieval method aiming at the problems of high computational complexities and low precision of global feature extraction algorithms. The establishment of the new retrieval method is on the basis of the SIFT and Harris (APISH) algorithm, and the salient region of interest points (SRIP) algorithm to satisfy users' interests in the specific targets of images. In the first place, by using the IPDSH and SRIP algorithms, we tested stable interest points and found salient regions. The interest points in the salient region were named as salient interest points. Secondary, we extracted the pseudo-Zernike moments of the salient interest points' neighborhood as the feature vectors. Finally, we calculated the similarities between query and database images. Finally, We conducted this experiment based on the Caltech-101 database. By studying the experiment, the results have shown that this new retrieval method can decrease the interference of unstable interest points in the regions of non-interests and improve the ratios of accuracy and recall.

Keywords: Image Processing, Image Retrieval, Interest Points, Salient Region, local distribution features.

1. Introduction

With the development of computer networks and multimedia technologies, there have been an increasing number of people who have searched for the information from the internet. This has been an indispensable part of people's daily lives. However, the rapid growth of data amount has restricted the progress of image retrieval [1-3]. In order to solve this problem, a growing number of scholars have invested in the researches of the C ontent-Based Image Retrieval (CBIR) systems. The CBIR system is a process of extracti ng query images' characteristics such as colors, shapes and textures. Meanwhile, it can ret rieve image information by comparing the images' similarities. Although the CBIR has achieved remarkable successes in some fields, the semantic gap has restricted its development. Therefore, finding accurate low-level visual feature has become a key factor of extracting implicit high-level semantic.

In respect to the researches of low-level visual features, the local characteristics methods for image retrievals have commonly been used. They include two categories which are segmentation based methods and interest points based methods [4]. Segmentation methods are the methods of using some image segmentation technologies to find the regions of significant objects. Nevertheless, the present segmentation technologies have very low accuracies to retreive images, so many scholars began to study the interest points based methods. For instance, Li Yangxi *et al.* [5] have proposed a query difficulty guided image retrieval system, which allowed him to apply ranking optimization approaches to predict the queries' ranking performances in terms of their difficulties. Another scientist Wang Xiaoyang *el al.* [6] have made an effort on questing a robust color image retrieval method by using visual interest point feature of significant bit-planes. He has proved that it was more effective to retrieve the noise-attacked images including blurring, sharpening and illumination by adopting this method. Moreover, Zhu Xianqiang *et al.* [7] have proposed a novel local saliency features by adopting interest extraction algorithms which could improve process efficiencies and fit to human visual perceptual characteristcs. At last, Glauco V. Pedrosa et al. [8] have defined the saliencies of a shape as the higher curvature points along the shape contour, and proposed a shape-based image retrieval technique using salience points to describe shapes, which is consists of a salience point detector, a salience representation using angular relative position and curvature value analyzed from a multi-scale perspective, and a matching algorithm considering local and global features to calculate the dissimilarity.

As a result, these scholars have approved that the interested point based method has changed the non-effective traditional ways of image retrievals. Specifically, the method has altered the way of directly transforming the interests points in the image matching algorithm to image retrieval. Meanwhile, the method considers more about the space distribution information of the interest points. It can improve the precisions to retrieve images. Nevertheless, interest points' detection algorithms can be influenced by unstable interest points in the regions of non-interests. Thus, the division of salient region has low

accuracy. Aiming at this problem, we have proposed a novel image retrieval method based on IPDSH and SRIP, which can perform better than some other popular algorithms. **Fig. 1** shows the structure of this image retrieval method:
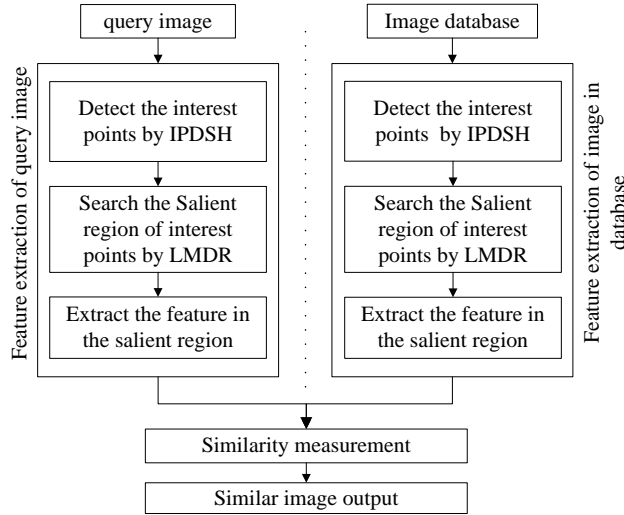

Fig. 1. Image retrieval structure

2. IPDSH

Interest points are usually defined as corners and there are some popular corner detection algorithms such as Harris and SIFT. In respect of the Harris algorithm, although this algorithm is effective in the aspects of high calculation speed and high computational precision, it is sensitive to scales which can lead to low accuracies of extractions. Specifically, the precision of corner extraction only depends on the threshold when response functions implement the non-maxima suppression and identify the local maxima. Meanwhile,the magnitude of threshold can determine the accuracy of detection. For example, corner information will lose as a result of a very large threshold and too small threshold can result in false corners [9]. To resolve this problem, some corner detection algorithms based on scale-invariant features were proposed. By saying that, the Scale-invariant feature transform (SIFT) algorithm is regarded as the most effective algorithm to slove the problem. Nevertheless, there are some disadvantages of the SIFT algorithm. For example, the computational complexity is high, calculation speed is low, and the interest points detected by the SIFT algorithm are not salient. Thus, we proposed the IPDSH algorithm based on the features of the SIFT and Harris (IPDSH) algorithm. The IPDSH algorithms can effectively improve the accuracies of interest points' detections. Specific procedures are as followings:

In the first place, since Koendetink [14] and Lindeberg [15] have proved that the Gaussian convolution kernel was the only constant linear scale transformation to achieve

volume. The image's $f(x, y)$ convolution can be realized with the two-dimensional Gaussian kernel $G(x, y, \sigma)$. Then, we can get the multi-scale spaces sequence S, as shown in equation (1).

$$S(x, y, \sigma) = G(x, y, \sigma) * f(x, y) \tag{1}$$

$G(x, y, \sigma)$ is defined as the equation (2). Included, $*$ represents the convolution, (x, y) represents the pixel position, and σ represents the scale space factor. If σ is getting smaller, the image smooth is getting less, and the corresponding scales will get smaller as well.

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \tag{2}$$

In order to improve the efficiency of stable interest points in the scale spaces, Lowe used the difference of Gaussian function (DoG) to get the convolution of $f(x, y)$, and obtain the maxima of the multi-scale spaces, as shown in equation (3).

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * f(x, y) = S(x, y, k\sigma) - S(x, y, \sigma) \tag{3}$$

where, $k = \sqrt{2}$ is a constant. Lindeberg found that DoG function was very similar to the normalized Gauss Laplasse function $\sigma^2 \nabla^2 G$, so the DoG function can be used to calculate the position of the extreme points in the scale spaces.

Futhermore, the descriptor of the SIFT interest points are based on the pixel gradients. Thus, in the local field of interest points, the larger changes of the pixel gradients can result in more unique descriptor vectors. At the same time, the matching of descritors will probably be more accurate. This gives us inspiration that if there are harris corners existing in the local field of the extreme points, the extreme points are more likely to be the stable interest points in the scale spaces.

The autocorrelation matrix is used to extract the features and describe the local structure of images. It is defined as the equation (4).

$$A(x, y, \delta_i, \delta_d) = \delta_d{}^2 G(\delta_i) * \begin{bmatrix} f_x^2(x, \delta_d) & f_x f_y(x, \delta_d) \\ f_x f_y(x, \delta_d) & f_y^2(x, \delta_d) \end{bmatrix} = \begin{bmatrix} \hat{f}_x^2 & \hat{f_x f_y} \\ \hat{f_x f_y} & \hat{f}_y^2 \end{bmatrix} \tag{4}$$

In which δ_i represents the integral, δ_d represents the differential scales, and $\hat{f}$ represents the Gaussian filtering of f, f_x and f_y represents the derivatives in the directions of x and y.

Suppose that the eigenvalues of A are λ_1 and λ_2, which represent the principal curvature of autocorrelation function, we can define the detection formula of different scale spaces as the equation (5).

$$C(x, y, \delta_i, \delta_d) = \det(A(x, y, \delta_i, \delta_d)) - \alpha \cdot trace^2(A(x, y, \delta_i, \delta_d))$$
$$= \lambda_1 \cdot \lambda_2 - \alpha \cdot (\lambda_1 + \lambda_2) \tag{5}$$

Using the equation (4) and (5), we can obatin the harris corners of scale spaces. Then, according to the magnitudes of λ_1 and λ_2, we can calculate C's local maxima. α is set to be a constant number within 0.04 and 0.06. After, we can analyze if the Harris corners exist in the local field of extreme points. If any Harris corner falls on the extreme points' neighborhood in the multi-scale spaces, we should retain this extreme point and regard it as the stable interest points. On the other hand, if any Harris corner does not exist, we should deleted this extreme point and regard it as the unstable interest point.
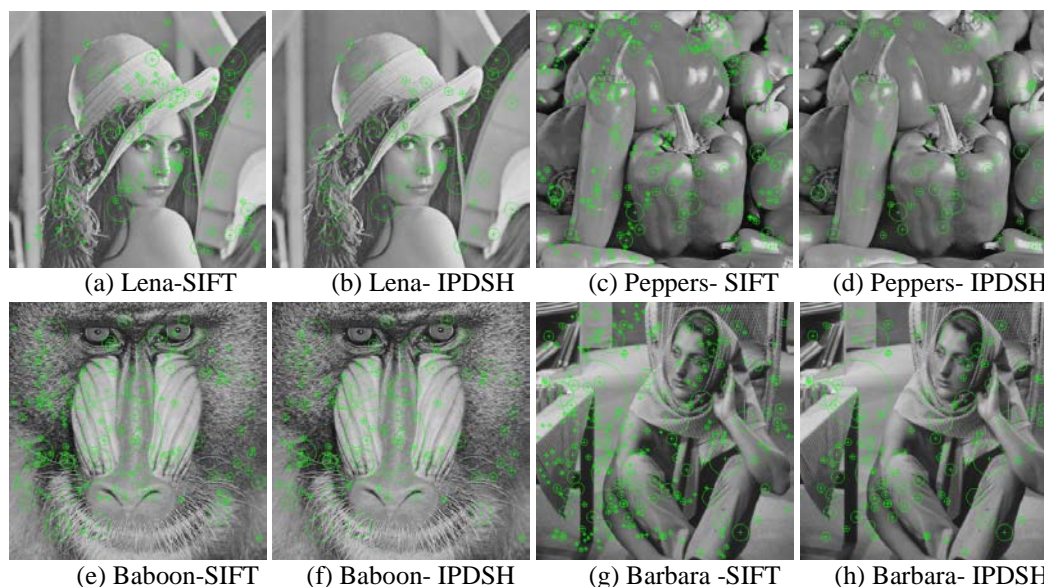

(a) Lena-SIFT   (b) Lena- IPDSH   (c) Peppers- SIFT   (d) Peppers- IPDSH

(e) Baboon-SIFT   (f) Baboon- IPDSH   (g) Barbara -SIFT   (h) Barbara- IPDSH

Fig. 2. Comparison of detecting results of IPDSH and SIFT

This method can effectively decrease the number of interest points and enhance the significance of interest points' set to improve the precision of image retrieval. **Fig. 2** shows an experiment result in comparing the IPDSH and the SIFT algorithms. It indicates the total number of interest points detected by IPDSH is lower than the SIFT algorithms because some unstable interest points have been removed.

3. SRIP

The vast majority algorithms of the CBIR are based on the global information, which describe images by using some global feature descriptors. The global features algorithms are simple, and less sensitive to translate or rotate. However, although the methods can satisfy some users' retrieval purposes, they still cannot describe the spatial differences of the image. As a result, the retrieval algorithms based on the global features have great limitations of satisfying the users' demands. Therefore, the CBIR systems are based on the local features should be analyzed in order to truely cope with the users' needs.

For instance, convex hull and annular division are two effective algorithms based on the local features. **Fig. 3** shows the feature regions of convex hull and annular division. Obviously, the accuracy of divisions has been influenced by the interest points located in the background. In general, most interest points of images are concentrated in the inner or edge of the objects. Some of these interest points are freely distributed in the background. When the total number of interest points is less, the free distribution points can disturb the division results.Thus, its negative effects should be considered in the feature of the extraction of image.
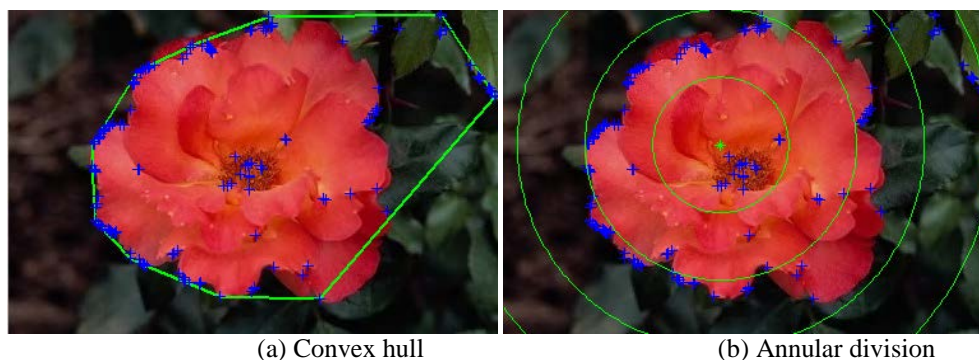

(a) Convex hull   (b) Annular division
Fig. 3. Division of feature region

By studying the convex hull and annular division above, these two alogrithms have negative effects on image extraction because of the disturbances of freely distributed interests ponits. Therefore, SRIP method which is also based on local features will be introduced.

In most time, users just want to find one or more specific targets of query images. Thus, it is significant to focus on the salient regions that users concern about. A salient region can express the main content of images well. To some extent, it can eliminate the interferences of redundant information, and be well accepted in the domain for image retrieval.

Furthermore, the salient region of images regards as a region that interest points' distribution density is obviously higher than the others. Higher density represents more particular information. Within the salient region, the interest points are called salient points. The interest points locate outside of the salient region need to be removed. After, the background information can be reduced and the feature extractions will have more accuracies. The detailed procedures of salient region division are as followings:

Suppose that the size of image is $m \times n$, the set of interest points P can be expressed as the equation (6). In which, $p(x, y)$ represents the interest points and (x, y) represents the coordinate.

$$P = \left\{ p(x, y) \middle| 1 \le x \le n, 1 \le y \le m \right\} \tag{6}$$

Let N represent the total number of interest points, and we define the centroid of interest points O as the equation (7).

$$O = (\overline{x}, \overline{y}) = \left(\frac{1}{N} \cdot \sum_{(x,y) \in N} x, \frac{1}{N} \cdot \sum_{(x,y) \in N} y \right) \tag{7}$$

Next, we calculate the distance between O and each interest points according to the equation (8), where, R_i is the distance from the $i-th$ interest points to the centroid O.

$$R_i = \left[(x_i - \overline{x})^2 + (y_i - \overline{y})^2 \right]^{1/2}, (x_i, y_i) \in P, i = 1, 2, \cdots N \tag{8}$$

Then, arranging R_i from large to small orders and setting the $k-th$ ($k \in N$) value to be radius that was recorded as R_c. We draw a circle C around the centroid, which is seen as the circle center. In **Fig. 4**, we determined an ergodic region according to the R_c, and iterated each pixel in this region. All pixels in the ergodic region can be seen as the center of circle, and then we will obtain some circles with R_c radius. In **Fig. 4**, A', B', C' and D' represent the four corners of the image. The blue region constitutes of A, B, C and D is the ergodic region. AA', BB', CC' and DD' are angle bisectors of the each corner of the image. The goal is to achieve to iterate all the pixels in the image by only iterating the least number of pixels. Moreover, if R_c is large enough, the ergodic region will become a line of EF.
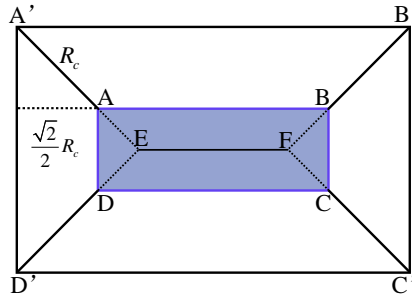

Fig. 4. Ergodic region

In the process of iteration, we recorded the coordinate of the center and the total number of interest points in the circle. After the comparison of each circle, we selected the one which the number of interest points was the largest. Then it was defined as the salient region. It was drawn on the image in green. Finally, we retained the salient interest points in the salient region and excluded the free distributed interest points, so as to obtain the feature extraction region. Additionally, the set of salient interest points is determined as the equation (9).

$$P' = \left\{ p(x', y') \middle| \sqrt{(x' - x_c)^2 + (y' - y_c)^2} \le R_c \right\} \tag{9}$$

where, x_c and y_c represent the circle center of the salient region. **Fig. 5** shows two examples of SRIP's extractions ($k = 14$). It can be seen that the free distributed points

have been effectively omitted. Plus, the feature-region extraction are more accurate than the results in **Fig. 3**.
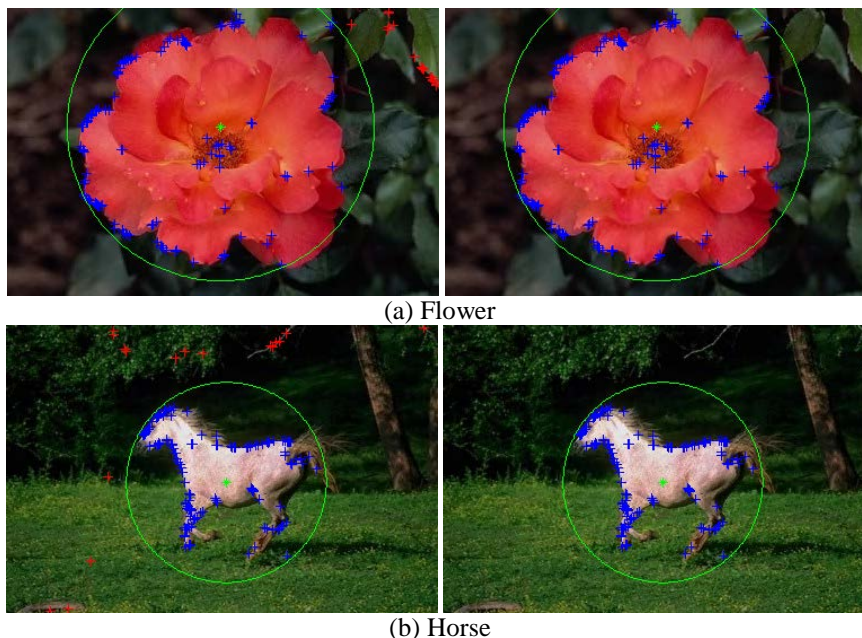

(a) Flower

(b) Horse
Fig. 5. Two examples of SRIP's extraction (k=14)

Additionally, by comparing the data, we found that the experiment could achieve the best results if $R_c = \lceil 0.9 \times N \rceil$, in which $\lceil \cdot \rceil$ rounds the elements to the nearest integers towards plus infinity, N represents the total number of stable interest points.

4. Pseudo-Zernike Moments

Pseudo-Zernike Moments [10] can be obtained from a group of basic functions $\{V_{nm}(x, y), x^2 + y^2 \leq 1\}$, which perform better than Zernike moment in anti-noise aspect. $V_{nm}(x, y)$ is defined as equation (10), which can satisfy the equation (11).

$$\begin{cases} V_{nm}(x, y) = V_{nm}(\rho, \theta) = R_{nm}(\rho) \exp(jm\theta) \\ R_{nm}(\rho) = \sum_{s=0}^{n-|m|} \dfrac{(-1)^s (2n+1-s)! \rho^{n-s}}{s!(n+|m|+1-s)!(n-|m|-s)!} \end{cases} \tag{10}$$

$$\begin{cases} \iint\limits_{x^2+y^2 \leq 1} V_{nm}^*(x, y)V_{pq}(x, y)dxdy = \dfrac{\pi}{n+1}\sigma_{np}\sigma_{mq} \\ \sigma_{np} = \begin{cases} 1 & n = p \\ 0 & n \neq p \end{cases}, \quad \sigma_{mq} = \begin{cases} 1 & m = q \\ 0 & m \neq q \end{cases} \end{cases} \tag{11}$$

where, n represents the non-negative integer, m is the integer ($|m| \leq n$), ρ is radius and θ is the angle of polar coordinates pixels. Then, we defined the $f(x, y)$'s (n, m) step pseudo-Zernike Moments as the following equation (12).

$$P_{nm} = \frac{n+1}{\pi} \iint\limits_{x^2+y^2 \leq 1} f(x, y)V_{nm}^*(x, y)dxdy = \frac{n+1}{\pi} \sum_{\rho \leq 1} \sum_{0 \leq \theta \leq 2\pi} f(\rho, \theta)V_{nm}^*(\rho, \theta)\rho$$

$$= \frac{n+1}{\pi} \sum_{\rho \leq 1} \sum_{0 \leq \theta \leq 2\pi} f(\rho, \theta)R_{nm}(\rho)\exp(-jm\theta)\rho \tag{12}$$

After confirming the interest points' location, we can build the origin of polar coordinates, calculate the pseudo-Zernike and convert unit circle to polar coordinates. The pixels locate outside of the polar coordinates will not be calculated.

5. Similarity Measurement

Suppose that Q represents the query image with E salient interest points, I represents the database image with F salient interest points, and N represents the total number of each salient interest points' Pseudo-Zernike Moments. Then, we can define the similarities of Pseudo-Zernike Moments S in salient interest points' neighborhood as the equation (13).

$$\begin{cases} S(Q, I) = 1/D(Q, I) \\ D(Q, I) = \left| \sum_{i=0}^{E} \sum_{j=0}^{N-1} \left\| P_Q(i) \right\| - \sum_{i=0}^{F} \sum_{j=0}^{N-1} \left\| P_I(i) \right\| \right| \end{cases} \tag{13}$$

where, $\left\| P_Q(i) \right\|$ represents the amplitude of the query images' $i-th$ moment, $\left\| P_I(i) \right\|$ is the same amplitude of the database images.

6. Experiments

The platform of "hardware: Core 2 Q8200 CPU, 2.33GHz, 4.0GB memory, OS: windows server 2003，software: Matlab7.1" and the Caltech-101 dataset have been selected to study this experiment. The Caltech-101 dataset totally consists of 9146 images, and splits to 101 different object categories as well as an additional background/clutter category. There are around 40 to 800 images per category, e.g., sunflower, butterfly, schooner, rhino, panda, lamp, cougar face, rooster, piano, watch, cup, lotus, barrel, ant, chairs and other images.

Fig. 6. Thumbnails from some categories of the Caltech-101 database

Fig. 6 shows the thumbnails of some categories.In this experiment, the precision is defined as $P_T = n / T$, in which T represents the total number of image outputs in the retrieval results and n represents the number of correct images. Some experiment results are shown in **Fig. 7**.

From the **Fig. 7**, there are 6 rows and each row has 10 images. The leftmost image of each row is query image, which were chosen from the categories shown in **Fig. 6**. Other images in **Fig. 7** in each row are the retrieval results which are arranged from large to small order according to the similarities. Furthermore, the images marked with red ellipses are the results in errors. In the first row, because there is no error result, the P_{10} of sunflower is 100%. In the second row, there are 3 error results, so the P_{10} of cougar face is 70%. Similarly, the P_{10} of the schooner, butterfly, panda and grand piano are respectively 70%, 80%, 60% and 60%. It is also important to note that the 3 error results in the third row are all ketches because it regarded ships as the same as schooners. If we deem the query image as ships, P_{10} will be valued 100%. In short, the 6 precisions of experiment results are over 50%, which can satisfy our needs.
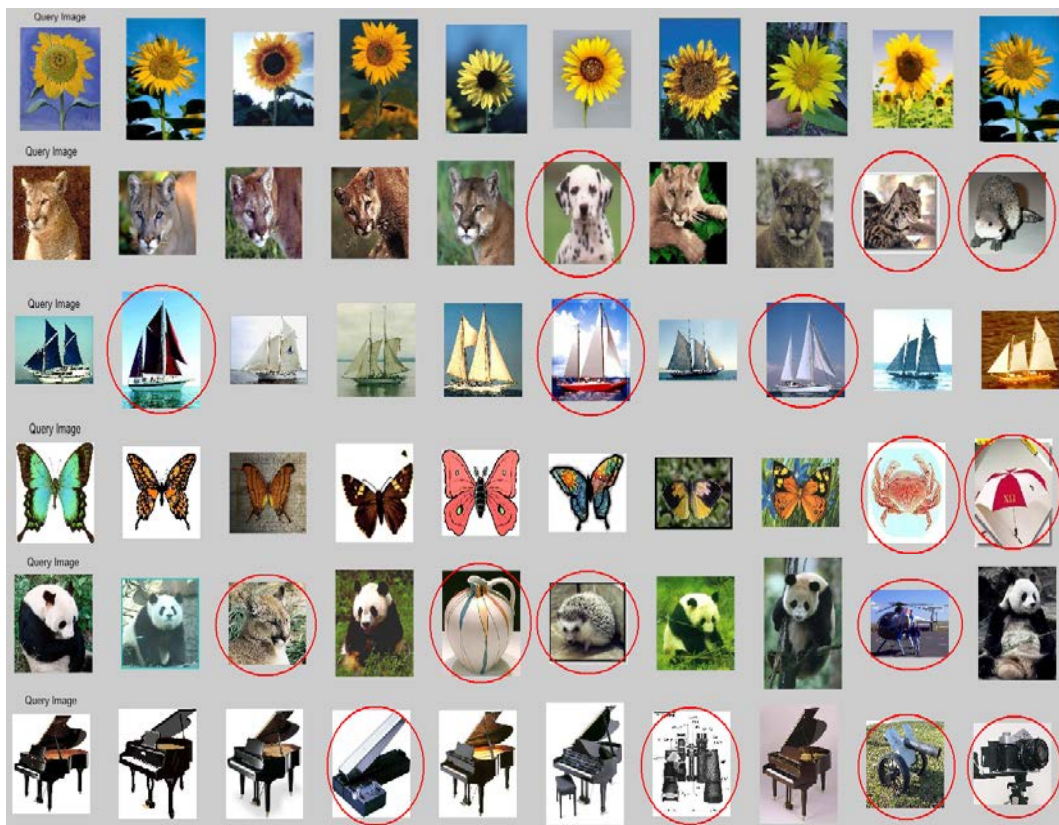

Fig. 7. Retrieval results

By comparing the corrected results and results in errors, it is obvious that the image shapes in errors are similar to the corrected image shapes. The reason behind this is that Pseudo-Zernike moments can describe these images well since it has space geometric invariance characteristis ofselecting, translating, and scaling.However, singly adopting this method can easily cause the distrubances of similar images. The advantage of this method is that the complexity is low and calculation speed can be fast. Nevertheless, if intergating with some algorithms based on more features such as colors and textures, the method we proposed can be improved.
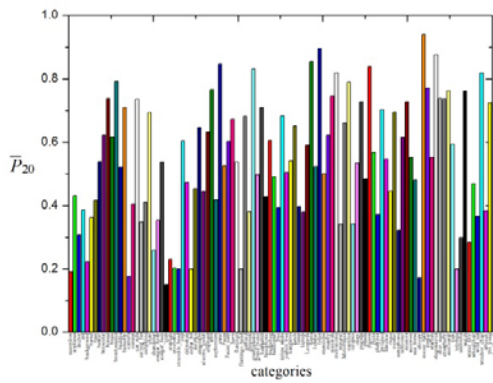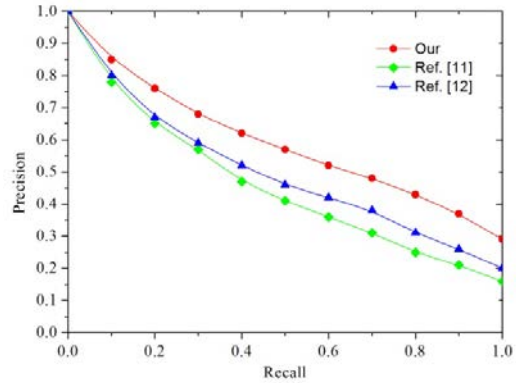

Fig. 8. $\overline{P}_{20}$ of 102 categories

Fig. 9. Comparison result of PR curves

Table 1. Precision of different methods

Categories	$\overline{P}_{10}$ / %			$\overline{P}_{20}$ / %			$\overline{P}_{30}$ / %		
	Our	Ref.[11]	Ref.[12]	Our	Ref.[11]	Ref.[12]	Our	Ref.[11]	Ref.[12]
sunflower	82.5	78.0	72.3	76.3	66.4	60.1	72.8	63.4	54.7
cougar_face	62.0	56.5	51.0	53.8	43.6	39.7	47.2	38.6	37.2
schooner	65.3	58.2	50.8	61.5	53.5	48.3	55.7	46.7	42.6
butterfly	76.5	75.3	68.5	70.9	64.0	61.6	64.1	60.4	56.8
panda	55.0	48.0	43.8	48.4	42.8	37.5	41.5	36.2	32.4
grand_piano	56.7	52.4	48.0	49.7	44.2	41.5	43.0	38.3	35.6
Average	66.3	61.4	55.7	60.1	52.4	48.1	54.1	47.3	43.2

In order to evaluate the performances of our algorithms further, we have randomly selected 10 query images in each category and calculated the average precisions $\overline{P}_{10}$, $\overline{P}_{20}$ and $\overline{P}_{30}$ of 6 categories mentioned above. The purpose is to compare the precision of both of the Chen's algorithms in Ref. [11] and Yu's algorithms in Ref. [12]. **Table 1** shows the precision of different methods. By looking at the Table1,the average precision of our method has increased 6.5 percent than Chen's algorithms and 11.2 percent than Yu's algorithms. Furthermore, **Fig. 8** shows the $\overline{P}_{20}$ of the 102 categories, which are listed alphabetically from left to right. In accordance with this figure, the soccer ball's precision is the highest, and the crab's precision is the lowest. Moreover, there are 60 categories' precisions are over 50 percent.
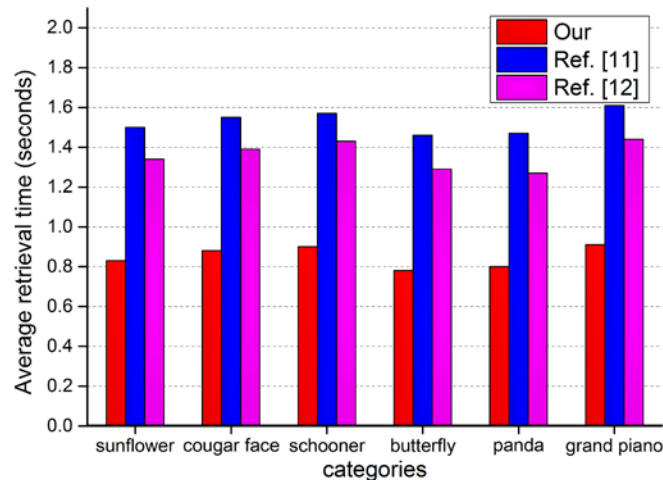
Fig. 10. Average retrieval time of 6 categories in the Caltech-101 database

Additionally, we have compared the Precision-Recall (PR) curves and the average retrieval time with the algorithms of Chen's and Yu's. **Fig. 9** shows the comparison result of PR curves. **Fig. 10** shows the average retrieval time of 6 categories in the Caltech-101 database. By reading this figure, our proposed method is superior than the other methods since it has effectively retrieved the precisions and the time complexities. Thus, the combination of IDPSH and SRIP can effectively reduce the amount of calculation, and accurately extract the images' visual features.

8. Conclusions

In this paper, we proposed an image retrieval method based on IPDSH and SRIP. The IPDSH algorithm can detect interest points more accurately after eliminating unstable interest points. On the other hand, the SRIP algorithm can reduce the interferences of some free distributed interest points in the images' backgrounds. As a result, it can make the feature extraction algorithm perform better. In addition, when confirming the salient regions, the Pseudo-Zernike moments will be extracted as the image features since they have the advantages of anti-noise, and rotational invariance. Finally, the retrieval results can output according to the similarities. To conclude, these experiment results have shown the proposed method we conducted can perform better rather than some other popular methods.

In the future, based upon other scholars' researches, we would like to extend our proposed methods to solve more challenging problems such as video searches, large data and mobile searches [13]. Furthermore, we are also going to make an efforts on improving detection speeds while guaranteeing accuracies.

References

[1] R. Datta, D. Joshi, J. Li, and J. Z. Wang, "Image retrieval: ideas, influences, and trends of the new age," *ACM Computing Suerveys*, vol. 40, no. 2, Article 5:1-59, April, 2008. Article (CrossRef Link).

[2] R. R. Ji, L. Y. Duan, J. Chen, L. X. Xie, H. X. Yao and W. Gao, "Learning to distribute vocabulary indexing for scalable visual search," *IEEE Transactions on MULTIMEDIA*, vol. 15, no. 1, pp. 153-166, January, 2013. Article (CrossRef Link).

[3] C. L. Yan, "Accurate image retrieval algorithm based on color and texture feature," *Journal of Multim*edia, vol. 8, no. 3, pp. 277-283, June, 2013. Article (CrossRef Link).

[4] X. Fu and J. X. Zeng, "A novel image retrieval method based on interest points matching and distribution," *Chinese Journal of Lasers*, vol. 37, no. 3, pp. 774-778, March , 2010. Article (CrossRef Link).

[5] Y. X. Li, B. Geng, Z. J. Zha, D. C. Tao, L. J. Yang and C. Xu, "Difficulty guided image retrieval using linear multiple feature embedding, " *IEEE Transactions on Multimedia,* vol. 14, no. 6, pp. 1618-1630, December, 2012. Article (CrossRef Link).

[6] X. Y. Wang, H. Y. Yang, Y. W. Li and F. Y. Yang, "Robust color image retrieval using visual interest point feature of significant bit-planes," *Digital Signal Processing*, vol. 23, no. 4, pp. 1136-1153, July, 2013. Article (CrossRef Link).

[7] X. Q. Zhu, J. C. Huang, Z. F. Shao and G. Q. Cheng, "A new approach for interesting local saliency features definition and its application to remote sensing imagery retrieval," *Geomatics and Information Science of Wuhan University*, vol. 38, no. 6, pp. 652-655, June, 2013. Article (CrossRef Link).

[8] G. V. Pedrosa, M. A. Batista and C. A. Z. Barcelos, "Image feature descriptor based on shape salience points, " *Neurocomputing*, vol. 120, no. 23, pp. 156-163, November, 2013. Article (CrossRef Link).

[9] Y. T. Wang, Y. Q. Chen, J. Li and B. M. Li, "The Harris corner detection method based on three scale invariance spaces," *International Journal of Computer Science Issues*. vol.9, no.6, pp. 18-22, November, 2012. Article (CrossRef Link).

[10] X. B. Dai, T. L. Liu, H. Z. Shu and L. M. Luo, "Pseudo-zernike moment invariants to blur degradation and their use in image recognition," *Lecture Notes in Computer Science*, Vol. 7551, pp. 90-97, 2013. Article (CrossRef Link).

[11] M. S. Chen, S. Y. Yang, Z. J. Zhao, P. Fu, Y. Sun, X. N. Li, Y. Sun and X. Y. Qi, "Multi-points diverse density learning algorithm and its application in image retrieval, " *Journal of Jilin University (Engineering and Technology Edition)*, vol. 41, no. 5, pp. 1456-1460, September, 2011. Article (CrossRef Link).

[12] J. Yu, Z. C. Qin, T. Wan and X. Zhang, "Feature integration analysis of bag-of-features model for image retrieval," *Neurocomputing*, vol. 120, pp. 355-364, November, 2013. Article (CrossRef Link).

[13] R. R. Ji, L. Y. Duan, J. Chen, H. X. Yao, J. S. Yuan, Y. Rui and W. Gao, "Location discriminative vocabulary coding for mobile landmark search," *Int J Comput Vision*, vol. 96, no. 3, pp. 290-314, February, 2012. Article (CrossRef Link)

[14] J.J. Koenderink. "The structure of images," *Biological Cybernetics*, vol. 50, pp. 363-396, 1984. Article (CrossRef Link).

[15] T. Lindeberg, "Detecting salient blob-like image structures and their scales with a scale-space primal sketch: a method for focus-of-attention," *International Journal of Computer Vision*, vol. 11, no. 3, pp. 283-318, 1993. Article (CrossRef Link).
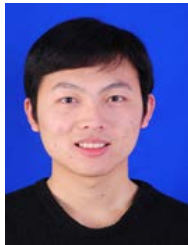
Xu Zhang received his B.S. degrees in control technology and instrument from Xidian University, Xi'an, and P.R. China in 2007. He is currently a Ph.D. student with the School of Areospace Science and Technology, Xidian University. His research interests include computer vision, pattern recognition, signal processing and image retrieval.

Baolong Guo received his B.S., M.S. and Ph.D. degrees from Xidian University in 1984, 1988 and 1995, respectively, all in communication and electronic system. From 1998 to 1999, he was a visiting scientist at Doshisha University, Japan. He is currently the director of the Institute of Intelligent Control & Image Engineering (ICIE) at Xidian University. His research interests include neural networks, pattern recognition, intelligent information processing, and image communication.

Yunyi Yan received his B.S. degree in automation, M.S. degree in pattern recognizaion andintelligent system, and Ph.D degree in electronic science and technology from Xidian University in 2001, 2004 and 2008 respectively in China. Since 2004, he has been with ICIE Institute of Areospace Science and Technology School of Xidian University, where he is currently an associated professor. Since 2013, he served as the director of Intelligent Detection Department in Xidian University. He worked as a visiting scholar in Montreal Neurological Institute of McGill University,Canada from 2010 to 2011. His research interests include system reliablity analysis, medical image processing, and swarm intelligence.

Wei Sun was born in 1980, Anhui province, China. He received his Bachelor degree in Measuring & Control Technology and Instrumentations, Master and Ph.D. degrees in Circuit and System from Xidian University, Xi'an, China, in 2002, 2005 and 2009, respectively. Since 2012, Dr. Wei Sun has been an associate Professor in School of Aerospace Science and Technology at Xidian University, Xi'an, China, His research interests include visual information perception, pattern recognition and embedded video system. He has published over 20 papers in the significant journal or conference and has been granted 3 patents.

Meng Yi (S'12) received the M.S. degree in Electrical Engineering from Northwestern Polytechnical University, Xi'an, China, in March 2008. Since 2009, he has been a Ph.D. of Electric circuit and systematic at Xidian University. Currently, he is a visiting doctoral candidate in Department of Electrical and Computer Engineering and Center for Automation Research, University of Maryland, College Park, maryland, USA. His research interests include computer vision, pattern recognition, signal processing and biometrics.