Available online at www.sciencedirect.com
SciVerse ScienceDirect
Procedia Engineering 29 (2012) 3371 – 3375
Procedia Engineering
www.elsevier.com/locate/procedia

2012 International Workshop on Information and Electronics Engineering (IWIEE)

Localized Image Retrieval Based on Interest Points

Meng Fanjie*, Guo Baolong, Wu Xianxiang

Institute of Intelligent Control & Image Engineering, Xidian University, Xi'an 710071, China

Abstract

This paper proposes a novel method for content-based image retrieval based on interest points. Interest points are detected from the scale and rotation normalized image. Then the normalized image is divided into a series of sector sub-regions with different area according to the distribution of interest points. With robustness to the image's rotation, scale and translation, local features of every sector sub-region are extracted to describe the image and make the similarity measure. In the relevant feedback phase, images are regarded as multi-instance (MI) bags, and the MI learning algorithm is employed to compute the target image feature. Finally, the similarity is recalculated. Experimental results show that our method can effectively describe the image, and obviously improve the average retrieval precision.

Keywords: content-based image retrieval;interest points; multi-instance learning; relevant feedback

1. Introduction

Traditional content-based image retrieval (CBIR) uses global features to retrieve images. As the large amount of computation for global features and the user is often interested in a certain object in the image, thus the use of local features for image feature description and retrieval have been proposed. Localized image retrieval [1] can be divided into image segmentation based methods [2-3] and interest points based methods [4-8]. The former obtain the contour or the region of the interested target via an image segmentation algorithm, and then extract the target feature to retrieval. However, image segmentation has been a key link to be solved. It is difficult to find a common segmentation algorithm to separate the target inextenso from the background. The latter locate the area which greatly changes in vision characteristics by interest points detection, and extract features around interest points by the local descriptors. Because

* Corresponding author. Tel.: 86-13891936755; fax: 86-029-88201777.
E-mail address: chinesemeng@sina.com.

doi:10.1016/j.proeng.2012.01.496

interest points detection algorithms have better versatility and stability compared with image segmentation algorithms, interest points based image retrieval attracted more and more research attention.

Wolf et al. [4] present the method by using Gabor features of interest points, which generates a series of histograms according to the maximum amplitude of different scales and directions to represent the image and retrieval. Zheng et al. [5] propose the retrieval method by combing invariants of interest points and the edge types histogram after interest points detection in every region of the image. Jian et al. [6] present the density based clustering algorithm to cluster interest points, and use colour moments and Gabor features of every class to describe the image and retrieval. Su et al. [7] calculate convex hulls of interest points recursively, and assign interest points into some buckets by a specific algorithm to form a color histogram for every bucket. Fu et al. [8] compute the best matching pairs of interest points and eliminate the dissimilar points by comparing the Euclidean distance of each interest point's local Zernike moments based on its local gray information.

These interest points based image retrieval methods above mentioned have changed the thinking of the early studies that transplanted image matching algorithms directly to the application of image retrieval, and more and more consider the spatial distribution of interest points. In order to further improve the accuracy of image retrieval, this paper proposes a new method for localized content-based image retrieval by using interest points and multi-instance (MI) learning [9]. The remainder of the paper is organized as follows: Section 2 describes the details of the proposed image retrieval method. Section 3 gives the experimental results. Finally, the conclusions are drawn.

2. Details of the proposed approach

2.1. Interest points detection

A good many of detectors for interest points have been proposed in the literature, with rotation and translation invariance Harris corner detector [10] is the most reliable one among these existed techniques. Because Harris interest points are sensitive to the image scale change, and in order to overcome the impact of the image rotation on image matching in the following steps, the image is firstly scale and rotation normalized [11] before Harris interest points detect.

2.2. Sector sub-region division and local feature description

This paper introduces a sector sub-region division method based on interest points, which can easily and effectively express the radiation distribution of interest points from the center to the surrounding. By using this regional division method the extracted feature will contain the spatial distribution information of interest points with the geometric invariant and achieve a more accurate description of the image. The process of the sector sub-region division method is described as follows.

Let (x', y') be an interest point of the normalized image $f(x, y)$, Ω be the set of interest points of $f(x, y)$, N be the number of interest points of $f(x, y)$.Let $O = (x', y')$ be the centroid of Ω .Set O as the centre of a circle, the distance between the farthest interest point and point O as radius R , generate the circle region. Divide the circle region into J sector sub-regions, each sector sub-region has $L = N / J$ interest points, where J is the integer which can make the L as an integer. In the following experiment, set $N = 240$, $J = 8$, $L = 30$. The division method shown in Fig. 1, in an anticlockwise direction from the initial radius R_0 on the right of the horizontal direction, the boundary radius R_1 of the first sector sub-region A_1 is determined by the Lth interest point. If there are interest points in radius R_0 , don't count in the sector sub-region A_1 . If there are any other interest points except the Lth in radius R_1 ,

count the nearest interest point from O in the first sector sub-region A_1, other interest points count in the second sector sub-region A_2. The rest can be done in the same manner.

Because the non-equal-area sector sub-region division method for the image is based on the same number of interest points in each sector sub-region, it can avoid the equal-area division may lead that the corresponding sector sub-regions of two different images have the same normalized colour histograms, and hence contribute to extract the image feature more accurately. The rotation normalization has been applied before the sector sub-region division, so it is able to overcome the matching error caused by the similar image content of two similar images in different degrees of rotation being divided into the sector sub-regions of different numbers, and hence can contribute to retrieve the query image more accurately.

For each sector sub-region, select the pixels in the δ neighborhoods of all interest points within it and calculate HSV space color histogram for them $h_k (1 \le \mathrm{k} \le \mathrm{M})$ [12]. In our experiment, we set $\delta = 5$. The spatial dispersion of interest points is used to represent the distribution density of interest points in each sector sub-region, which in the kth sector sub-region is calculated as follows:

$$D_k = \frac{1}{\mathrm{R}} \sqrt{\frac{1}{L} \sum_{(x',y') \in \Omega_k} (x' - \overline{x_k'})^2 + (y - \overline{y_k'})^2} \tag{1}$$

where Ω_k is the set of interest points in the kth sector sub-region, $(\overline{x_k'}, \overline{y_k'})$ is the centroid of Ω_k.

2.3. Similarity measure

Let Q and I be the query image and an image in the database. The similarity between the two images is given by

$$S = \omega_c S_c(Q,I) + \omega_d S_d(Q,I) \tag{2}$$

where ω_c and ω_d are two weight coefficients, $\omega_c + \omega_d = 1$. In our experiment, the two coefficients are regarded as 0.5. $S_c(Q,I)$ and $S_d(Q,I)$ represent the color histogram similarity and the spatial dispersion similarity between two images respectively. We employ histogram intersection to compute $S_c(Q,I)$. $S_d(Q,I)$ using Gaussian function is defined as

$$S_d(Q,I) = \frac{1}{J} \sum_{k=1}^{\mathrm{J}} \exp\left[-(D_k(Q) - D_k(I))^2\right] \tag{3}$$

2.4. Multi-instance learning

We take features of each sector sub-region corresponding to the concepts of instance in MI learning. Features of each sector sub-region is represented by nine colour moments in HSV colour space computed for pixels in the δ neighborhoods of all interest points in it, and one spatial dispersion of all interest points in it. That means each instance is taken as a point in the 10-dimensional vector space. From the first round of retrieval results, select the best front five positive images and five negative images as the positive and negative bags to constitute the training set, and apply EM-DD algorithm [13] to obtain the user desired target image feature q. The steps are as follows:

Step 1: Each instance in every positive bag in the training set is taken as the initial assumption target image feature q'; Step 2: Select the instance ε_e which is closest to q' for every bag in the training set:

$$\varepsilon_e = \arg\max_{B_{ef} \in B_e} \exp\left(-\sum_{z=1}^{10} (B_{efz} - q'_z)^2\right) \tag{4}$$

where B_{ef} denotes the instance f of the bag B_e, B_{efz} denotes the component z of the instance f of the bag B_e, q'_z denotes the component z of the assumption target image feature q';Step 3: Estimate a new target image feature q'' by the following formula:

$$q'' = \arg\max_{q'} \prod_{e} \Pr(q' | \varepsilon_e) \quad (5)$$

where, $\Pr(q' | \varepsilon_e) = 1 - |\gamma_e - \Pr(\varepsilon_e \in q')|$, γ_e is the label of the bag B_e, if B_e is positive, $\gamma_e = 1$, otherwise, $\gamma_e = 0$, $\Pr(\varepsilon_e \in q') = \max_{B_{ef} \in B_e} \exp(-\sum_{z=1}^{10} (B_{efz} - q'_z)^2)$; Step 4: Let $q' = q''$, repeat steps 2 and step3 until the algorithm convergences to obtain the target image feature $q = q'$.

With the target image feature q, calculate the distance between the target image feature and the most similar instance in every unlabeled image bag as the probability of the image containing the target image feature:

$$S_M = \max_{f} \left\{ \exp(-\sum_{z} (W_{fz} - q_z)^2) \right\} \quad (6)$$

where S_M denotes the probability of the image W containing the target image feature, W_{fz} denotes the component z of the instance f in the image W, q_z denotes the component z of the target image feature q. Combine the two similarities S and S_M to compute a new similarity S':

$$S' = S + S_M \quad (7)$$

3. Experiment Results

Image database used in the experiment is the test set of Simplicity system, which contains 1000 images collected from the Corel database. These images belong to 10 categories, each category includes 100 images. In order to verify the performance of our method, we also implement interest points based methods [7, 8] under the same software and hardware conditions. Retrieval precision is used as the image retrieval performance evaluation criterion. The precision is computed by:

$$P_T = n / T \quad (8)$$

where T is the number of images returned by the retrieval system, and n is the number of output images in the same category with the query image.

Taking out 20 images randomly as query images from each category, we compute the average precision for each category. Comparison results are illustrated in Fig. 2. Fig.2 (a)-(c) is the comparison of the precision for 10, 20, 30 images returned between our proposed method without MI learning and methods [7, 8] without relevance feedback; Fig.2 (d)-(f) is the comparison of the precision for 10, 20, 30 images returned between our proposed method with MI learning and methods [7] with relevance feedback. The X label 1-10 is corresponding to the category of Africa people and villages, beach, buildings, buses, dinosaurs, elephants, flowers, horses, mountains and glaciers, food. From Fig. 2 (a)-(c), we can see that the average precision of our proposed method is slightly better than the method [7] and is significantly better than the method [8] for all categories of images. From Fig. 2 (d)-(f), we can see after the introduction of MI learning, the average precision of our method is obviously better than the method [7] for all categories of images, and our method improves the average retrieval precision over 8.03 percent.

4. Conclusions

In this paper, we present a localized content-based image retrieve method by using interest points. The image is divided into a series of sector sub-region with different area according to the distribution of interest points, and local features with the spatial distribution of interest points are extracted, which is robust to the image's rotation, scale and translation. By the introduction of MI learning, the method further improves the accuracy of image retrieval. The future work is to study the application of generalized MI learning for interest points based image retrieval.

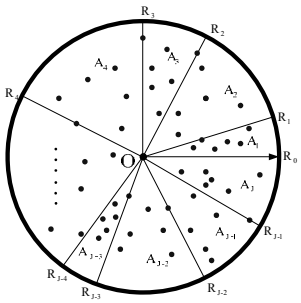

Fig. 1: Sector sub-region division

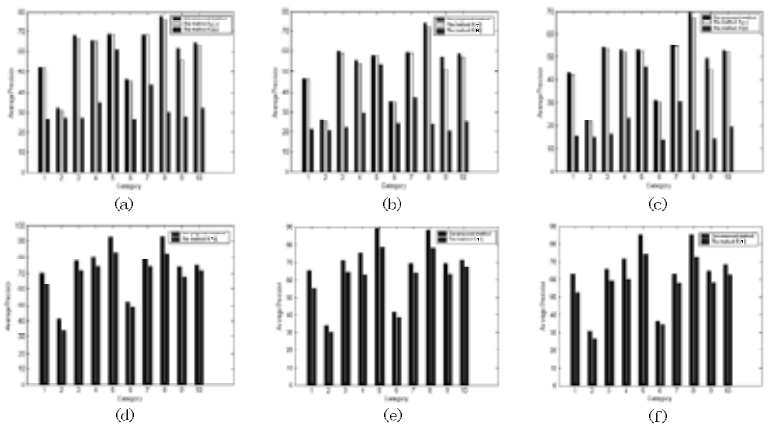

Fig. 2: Comparison results of different methods

Acknowledgements

This work is supported by the National Natural Science Foundation of China under Grants 61105066 and the Fundamental Research Funds for the Central Universities under Grants K50510040007.

References

[1] Rouhollah Rahmani, Sally A. Goldman, Hui Zhang, Sharath R. Cholleti, Jason E. Fritts. Localized content based image retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence.* 2008, 30(11): 1902-1912.

[2] Zeng Zhi Yong, Liu Shi Gang. A novel region-based image retrieval algorithm using hybrid feature. *WRI World Congress on Computer Science and Information Engineering*. Los Angeles, CA: IEEE Computer Society, 2009:416-420.

[3] Chiang Cheng-Chieh, Hung Yi-Ping, Yang Hsuan, Lee Greg C. Region-based image retrieval using color-size features of watershed regions. *Journal of Visual Communication and Image Representation.* 2009, 20(3):167-177.

[4] Christian Wolf, Jean-Michel Jolion, Walter Kropatsch, Horst Bischof. Content based image retrieval using interest points and texture features. *Proc. of 15th International Conference on Pattern Recognition*. Barcelona: IAPR, 2000: 234-237.

[5] Zheng Xia, Zhou MingQuan, Wang XingCe. Interest point based medical image retrieval. *Lecture Notes in Computer Science.* Beijing: Springer Verlag, 2008:118-124.

[6] Muwei Jian, Shi Chen. Image retrieval based on clustering of salient points. *Proc. of 2008 2nd International Symposium on Intelligent Information Technology Application*. Shanghai: Inst. of Elec. and Elec. Eng. Computer Society, 2008: 347-351.

[7] Su Xiao-Hong, Ding Jin, Ma Pei-Jun. Image retrieval by convex hulls of interest points and SVM-based weighted feedback.*Chinese Journal of Computers*. 2009,32(11):2221-2228.

[8] Fu Xiang, Zeng Jiexian. A novel image retrieval method based on interest points matching and distribution. *Chinese Journal of Lasers*. 2010, 37(3):774-778.

[9] Dietterich T G, Lathrop R H, Lozano-Pérez T. Solving the multiple-instance problem with axis-parallel rectangles. *Artificial Intelligence.* 1997, 89(1-2): 31-71.

[10] C. Harris, M. Stephens. A combined corner and edge detector. Proceedings of the Fourth Alvey Vision Conference, Manchester, 1988: 147-151.

[11] S.C. Pei, C.N. Lin. Image normalization for pattern recognition. *Image and Vision Computing*. 1995, 13(10): 711−723.

[12] Wu Jianhua, Wei Zhaorong, Chang Youli. Color and texture feature for content based image retrieval. *International Journal of Digital Content Technology and its Applications.* 2010, 4(3):43-49.

[13] Zhang Q, Goldman S A. EM-DD: an improved multiple-instance learning technique. *Advances in Neural Information Processing Systems 14.*Cambridge: CA: MIT Press, 2002:1073-1080.