

Image Retrieval using Harris Corners and Histogram of Oriented Gradients

K.Velmurugan
Dept. of Computer Applications
SNR Sons College, Coimbatore
Tamil Nadu ,India

Lt. Dr. S.Santhosh Baboo
PG & Research Dept. of Computer Applications
D G Vaishnav College, Chennai
Tamil Nadu, India

ABSTRACT

Content based image retrieval is the technique to retrieve similar images from a database that are visually similar to a given query image. It is an active and emerging research field in computer vision. In our proposed system, the Interest points based Histogram of Oriented Gradients (HOG) feature descriptor is used to retrieve the relevant images from the database. The dimensionality of the HOG feature vector is reduced by Principle Component analysis (PCA). To improve the retrieval accuracy of the system the Colour Moments along with HOG feature descriptor are used in this system. The Interest points are detected using the Harris-corner detector in order to extract the image features. The KD-tree is used for matching and indexing the features of the query image with the database images.

Keywords: Content-based image retrieval, interest points, HOG, KD-tree.

1. INTRODUCTION

Content-based image retrieval is emerging as an important research area with application to digital libraries and multimedia databases. It retrieves the similar images from a large collection according to the similarity between features extracted from the query image and candidate images based on the low-level features such as colour, texture and shape. The features are automatically extracted from the images themselves. It is an alternative to the conventional text-based image retrieval systems. The most CBIR systems use the querying images by examples (Query-By-Example) approach to search for the most similar images to the given example image among a number of candidate images. Some of the CBIR systems are QBIC (Query by Image Content), Virage, Pichunter, VisualSEEK, Chabot, Excalibur, Photobook, Jacob, Digital Library Project [1].

In general, image features can be either local or global [2]. The global features describe the visual content of the entire image. The retrieval systems based on global features cannot represent all the characteristics of the image. Therefore, the global features are not suitable for tasks like partial image matching or searching for images that contain the same object or same scene with different viewpoints. In order to avoid using global features, the interest points detectors were introduced to represent the local features of images in image retrieval systems[3,4]. The interest points are the salient image patches that contain rich local information about an image. Many

algorithms have been developed for the purpose of detecting and extracting the interest points like Harris[5], Hessian[6] , Scale invariant [7] , affine-invariant [8] and Difference-of- Gaussians (DOG) [9]. In this work the fast, robust and rotation invariant, Harris detector is used to detect the interest points. Local features have been initially proposed to solve problems in computer vision applications. They are very robust to occlusions, cropping and geometric transformations. In recent years, interest point detectors and descriptors are employed in many CBIR systems.

2. PROPOSED SYSTEM

In this paper, the interest points are used to denote the local features of an image. The 16 x 16 pixel image patch (region) is selected around the interest points to extract features. For each region, the HOG and colour features are extracted and form as the feature vector of the region. The similarity between a pair of images is determined by the number of matches found between their feature vectors. In this work, KD-Tree is used to index and match the similarity between the feature vectors of images and a voting algorithm is used to count the number of matches of feature vectors of query image with the database images. Finally the images having the maximum voting are ranked and extracted from the database. The Precision and Recall measure the performance of the system.. The Figure 1 has shown the proposed system of the CBIR.

The paper is organized as follows. Section 3 presents the processing steps of interest points detection using the Harris detectors, in section 4 the feature extraction algorithms is presented, the dimensionality reduction algorithm is explained in section 5, in section 6 the KD-tree is explained for matching the features of images, Section 7 demonstrates the experimental results and finally the paper concludes in Section 8.

3. INTEREST POINT DETECTION

In this work, the CBIR system retrieves the similar images using the interest points. Different Interest points detector have been proposed and used based on the field of applications. The fast, robust and rotation invariant, Harris detector [4] is widely used in many computer vision applications which uses the autocorrelation function to determine locations where the change of signal in one or two directions. A matrix related to the auto-correlation function is computed:

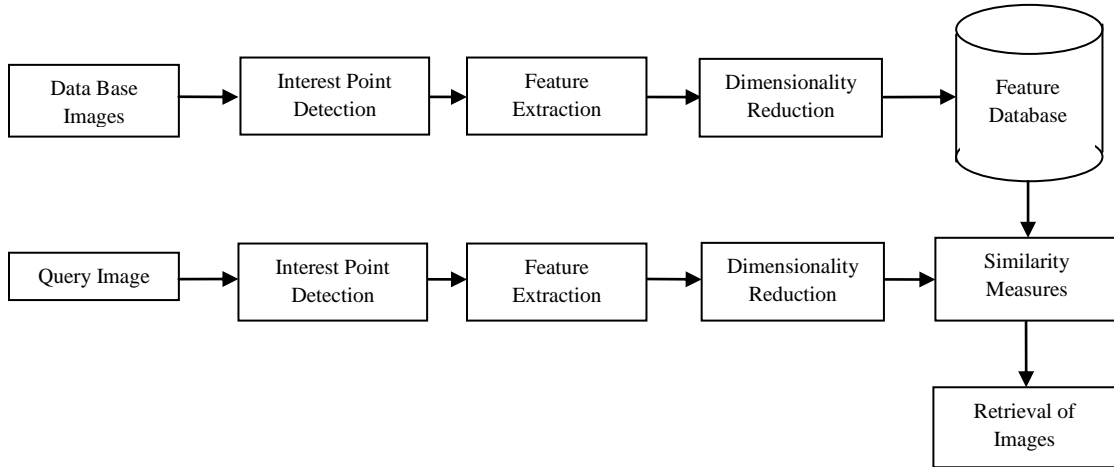


Figure 1: The Proposed CBIR System

$$C(x, \sigma_I, \sigma_D) = \sigma_D^2 G(x, \sigma_I) * \begin{pmatrix} L_x^2(x, \sigma_D) & L_x L_y(x, \sigma_D) \\ L_x L_y(x, \sigma_D) & L_y^2(x, \sigma_D) \end{pmatrix} \quad (1)$$

where σ_D is the derivation scale, σ_I is the integration scale, G is the Gaussian and L is the image smoothed by a Gaussian kernel. This matrix has two Eigen values that are the principal curvatures of the auto-correlation function. When the two eigenvectors are very small then there is no structure exists. If one is large and another one is small, there is an edge like structure. If both of them are very large and distinct, there is a corner like structure. Edges and interest points can be computed based on:

$$\det(C) - \alpha \cdot \text{trace}^2(C) < T_E \quad (2)$$

and

$$\det(C) - \alpha \cdot \text{trace}^2(C) > T_C \quad (3)$$

Edges are computed based on equation (2), where α is the coefficient of the Harris function and T_E is the threshold of the Harris function ($T_E < 0$). The edge detection is carried out at the first scale. Interest points can be detected by using eq. (3), T_C is the threshold for interest points ($T_C > 0$). The Figure 2 shows the detected interest points of a cat image.

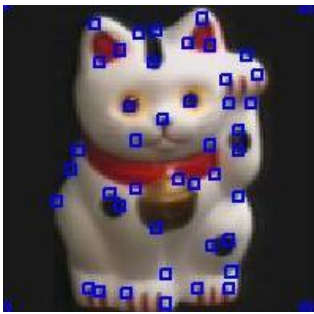


Figure 2: Harris Corners of a Cat

4. FEATURE EXTRACTION

The Feature extraction is the main core for the CBIR systems. The extracted features should be well separated in the feature space to produce effective discrimination between images. In this work the texture feature is extracted using HOG descriptor and the colour information is extracted using colour moments around the interest points.

4.1 Histogram of Oriented Gradients (HOG) Feature

Histogram of Oriented Gradients (HOG) [10] feature descriptor is used in this work to extract the features. It is very effective to represent objects and is widely used in human and face detection. The first step is detecting all interest points of the image using the Harris detector. This operator is based on the auto-correlation matrix that describes the local structure of the image. Then compute the Gradient Orientation Histogram around the 16 x 16 pixel region of each interest points. First, the region is divided in 4 X 4 sub-region, For each sub-region the 8-bin gradient orientation $h(k)$, $k = 0$ to 7 are calculated which forms a feature vector of size 128 dimension (4 x 4 x 8). The gradient oriented histogram is computed as follows:

$$h(k) = \sum_{a_{ij} \in a_k} m_{ij} \quad (4)$$

$$m_{ij} = \sqrt{dx_{ij}^2 + dy_{ij}^2} \quad (5)$$

$$d_{ij} = \arctan \frac{dy_{ij}}{dx_{ij}} - D \quad (6)$$

$$dx_{ij} = I_{ij} - I_{i+1,j} \quad (7)$$

$$dy_{ij} = I_{ij} - I_{i,j+1} \quad (8)$$

$$D = \arctan \frac{\sum_{ij} dy_{ij}}{\sum_{ij} dx_{ij}} \quad (9)$$

where I_{ij} is the i,j pixel value of each sub-region, m_{ij} is the gradient magnitude of the pixel i,j , d_{ij} is the gradient direction at pixel i,j , $h(k)$ is the k^{th} dimension $h(k)$ of the gradient histogram represents the total intensity of the pixel gradient whose direction lies in the k^{th} direction bin d_k , $k = 0$ to 7 . The direction bins are defined by the relative angle to the dominant gradient direction D of the image region. Finally, combing all the Gradient orientation Histogram of the interest point's area together to form a feature vector of size 128-dimension.

4.2 Colour Feature

The HOG gives only texture information to extract the color features around the region of each interest points the Color Moments [11] are used. The first order moment (mean) and second order(variance) are calculated for a 10x10 region around the interest point for the RGB channel. The value of the i -th color channel at the j -th image pixel is p_{ij} . The index entries related to this color channel are calculated by:

$$E_i = \frac{1}{N} \sum_{j=1}^N P_{ij} \quad (10)$$

$$\sigma_i = \left[\frac{1}{N} \sum_{j=1}^N (P_{ij} - E_i)^2 \right]^{\frac{1}{2}} \quad (11)$$

where N is the number of pixels in the image patch.

5. DIMENSIONALITY REDUCTION

Principal Component Analysis (PCA) [12] is applied to reduce the dimensionality of the feature vectors. PCA is a multi-variate statistical analysis method used for feature reduction. The new features are linear transformations of the original variables and obtained such that they are uncorrelated or orthogonal to each other. The HOG feature vector is reduced from 128 to 64 dimensions. The color feature is added after the feature vector dimension is reduced. So the final feature vector gets a size of 70 dimensions.

6. INDEXING AND MATCHING

In our CBIR system the KD-tree [13] algorithm is used to match the features of the query image with those of the database images. The KD-tree with the Best Bin First(BBF) search algorithm is used for indexing and matching the HOG features. The KD-tree is a kind of binary tree in which each node chooses a dimension from the space of the features being classified: all features with values less or equal to the node in that particular dimension will be put in the left sub-tree; the other nodes will be put in the right sub-tree and thus recursively. The BBF algorithm uses a priority search order to traverse the KD-tree so that bins in feature space are searched in the order of their closest distance from the query. The k-approximate and reasonable nearest matches can be returned with low cost by cutting off further search after a specific number of the nearest bins have been explored. The Voting scheme algorithm is used to rank and retrieved the matched images. To check the performance of proposed technique the precision and recall is used. The standard definitions of these two measures are given by following equations.

$$Precision = \frac{\text{No. of relevant images retrieved}}{\text{Total no. of images retrieved}} \quad (12)$$

$$Recall = \frac{\text{No. of relevant images retrieved}}{\text{Total no. of relevant images in the database}} \quad (13)$$

The matching of feature vectors of two images is shown in Figure 3.

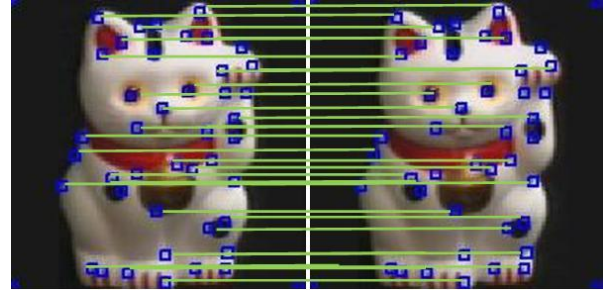


Figure 3: Matching of Feature points

7. EXPERIMENTS AND RESULTS

The image retrieval system based on Harris corners with HOG feature tested on COIL-100 object database [14]. The COIL-100 is a popular image database for benchmark which contains 72 views for 100 objects acquired by rotating the object under study about the vertical axis. In Figure 4, shows sample views for the each of the objects in the database. Our database consists of total 720 images of size 128x128. There are 10 different categories consisting of 72 images in each category. To test this system, five images are selected randomly from each category as query images. The HOG feature and colour moments are extracted for all images of the database which forms the feature database in off-line. The size of the HOG feature vector is reduced to 64 dimensions. The color information is added to the reduced HOG features. Now the size of the feature vector is 70(64-d HOG + 3 x 2 First order and second order colour moments of RGB channel). The fast and multidimensional KD-tree data structure is used to compare the features of the query image with the data base images. The retrieved result of coffee cup object query image is shown in Figure 5.



Figure 4: Sample database

Table 1: Retrieved images using HOG feature descriptor and Colour Moments

Object Category/Threshold Value	Object 1	Object 2	Object 3	Object 4	Object 5	Object 6	Object 7	Object 8	Object 9	Object 10	
		0.7	0.35	0.35	4 0.42	0.69	0.58	0.65	0.33	0.52	0.49
No. of relevant images retrieved	Qimage1	72	66	72	62	64	50	66	62	36	42
	Qimage2	71	69	72	61	62	49	72	52	39	40
	Qimage3	71	72	72	62	63	45	69	60	41	44
	Qimage4	72	68	72	63	64	52	69	53	40	39
	Qimage5	71	69	72	62	61	52	69	66	37	41
Average	71.4	68.8	72	62	62.8	49.6	69	58.6	38.6	41.2	
% Precision/Recall	99.1	95.5	100	86.1	87.2	68.8	95.8	81.3	53.6	57.2	

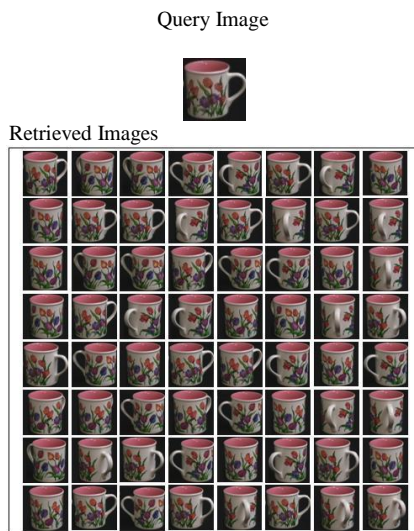


Figure 5: Result of Coffee Cup Object query image

8. CONCLUSION

Content-based image retrieval is currently a very important and active research in the field of multimedia databases. Since the explosive growth of image data in the large image archives need a more précised retrieval techniques to find the similar images. In this paper the interest points based Histogram of Oriented Gradients feature descriptor is used to retrieve the similar images from the database. The Harris detector is used to detect the interest points in the images. The HOG descriptor is used to extract the features around the interest points. The dimension of the HOG feature vector is reduced by the PCA and the colour information is added to the reduced HOG feature to improve the retrieval accuracy of the system. The proposed system gives 82.46% of Average Precision, for 10 categories of 720 images. This proposed system is more suitable for such applications as searching for images that contain the same object or the same scene with different viewpoints.

9. REFERENCES

- [1] R. C. Veltkamp and M. Tanase, "Content-based image retrieval systems:A survey," Department of Computing Science, Utrecht University, Tech.Rep. UU-CS-2000-34, 2000.
- [2] A. Halawani, A. Teynor, L. Setia, G. Brunner, and H. Burkhardt, "Fundamentals and Applications of Image Retrieval: An Overview", presented at Datenbank-Spektrum, 2006, pp.14-23.
- [3] C. Schmid and R. Mohr. Local grayvalue invariants for image retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(5):530.534, 1997.
- [4] T.Tuytelaars and L.Van Gool, "Content-based Image Retrieval based on Local Affinely Invariant Regions".3rd Int. Conf. on Visual Information Systems, Visual99, Amsterdam, The Netharlands, 2-4 June 1999,pp.493-500.
- [5] Harris and M.J. Stephens. "A combined corner and edge detector". In *Proc. Alvey Vision Conference*, pp. 147-152, 1988.
- [6] K. Mikolajczyk. *Interest point detection invariant to affine transformations*. PhD thesis. Institut National Polytechnique de Grenoble. 2002.
- [7] K. Mikolajczyk and C. Schmid. "An affine invariant interest point detector" in *Proc. of the 7th European Conf. on Computer Vision*. Copenhagen, Denmark, vol. I, pp. 128–142. 2002.
- [8] D. Lowe. "Object Recognition from Local Scale-Invariant Features", in *Proc. of the 7th Int. Conf. on Computer Vision*. vol. 2, p. 1150. September 20–25, 1999. IEEE, 1999.
- [9] K. Mikolajczyk and C. Schmid (2001), Indexing based on scale invariant interest points., *International Conference on Computer Vision*, pp. 525-531.
- [10] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In *CVPR (1)*, pages 886–893, 2005.

- [11] M. Stricker, and M. Orengo, "Similarity of color images", *SPIE Storage and Retrieval for Image and Video Databases III*, vol. 2185, pp.381-392, Feb. 1995.
- [12] Y. Ke and R. Sukthankar (2004), .PCA-SIFT: A more distinctive representation for a local image descriptors., Techn. Rep., *INTEL Research*.
- [13] Jacob E. Goodman, Joseph O'Rourke and Piotr Indyk (Ed.) (2004). "Chapter 39 : Nearest neighbours in high-dimensional spaces". *Handbook of Discrete and Computational Geometry (2nd ed.)*. CRC Press.
- [14] S.A. Nene, S.K. Nayar and H. Murase. "Columbia Object Image Library (COIL-100)", TR CUCS-006-96, Dept. Comp. Sc., Columbia University, 1996.
- [15] D. Lowe (2004), .Distinctive image features from scale-invariant keypoints., to appear in the *International Journal of Computer Vision*.
- [16] K. Mikolajczk and C. Schmid (2003), .A performance evaluation of local descriptors., *IEEE Conference on Computer Vision and Pattern Recognition*.
- [17] J.R. Smith and S. Chang. "Single color extraction and image query", in *Proc. of IEEE Int'l Conf. on Image Processing*, pp. 528-531, 1995.
- [18] L. Huston Y. Ke, R. Sukthankar, "Efficient near-duplicate detection and sub-image retrieval," in Proc. of ACM Multimedia (MM'04), 2004, 2004.
- [19] R. Datta, D. Joshi, J. Li, and J. Z. Wang, "Image retrieval: Ideas, influences, and trends of the new age," *ACM Comput. Surv.*, vol. 40, no. 2, pp. 1–60, 2008.
- [20] W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 12, pp.1349–1380, 2000.
- [21] C. Wolf, W. Kropatsch, H. Bischof, and J.-M. Jolion. Content based image retrieval using interest points and texture features. *International Conference on Pattern Recognition*, 04:4234, 2000.
- [22] H. Zhang, R. Rahmani, S. R. Cholleti, and S. A. Goldman. Local image representations using pruned salient points with applications to CBIR. In *Proceedings of the 14th Annual ACM International Conference on Multimedia*, 2006.
- [23] T. Pavlidis, Limitations of CBIR, In ICPR ,2008.
- [24] Xu Wangming Wu Jin Liu Xinhai Zhu Lei Shi Gang, "Application of Image SIFT Features to the Context of CBIR", *Int. Conference on computer Science and software engineering*, Wuhan, Hubei Volume:4 page(s):552-555, ISBN: 978-0-7695-33360.
- [25] Nguyen Duc Anh, Pham The Bao, Bui Ngoc Nam, Nguyen Huy Hoang, "A new CBIR system using Sift combined with neural network and graph-based segmentation", *ACIIDS'10 Proceedings of the Second international conference on Intelligent information and database systems: Part I* Springer-Verlag Berlin, Heidelberg ©2010, ISBN:3-642-12144-6 978-3-642-12144-9.
- [26] Zhang, R., and Zhang, Z, "FAST: toward more effective and efficient image retrieval", *ACM Multimedia Systems*, 10(6), 529-543, 2005..
- [27] Horster, E., Lienhart, R., and Slaney, M. 2007. "Image retrieval on large-scale image databases", In *Proceedings of the 6th ACM International Conference on Image and Video Retrieval* (Amsterdam, Netherlands), 17-24.