International Journal of Research
Available at https://edupediapublications.org/journals
e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 04 Issue-17
December 2017

Comparative Study of CBIR Systems Based on PCA and SIFTS Feature Extraction Methods

B. Shravan Reddy , M. Narsing Yadav , R. Raja Kishore
(PG Scholar)[1] (Professor, Ph.D)[2] (M.Tech, Assistant Professor)[3]
Malla Reddy Institute of Engineering and Technology, Medchal, Secunderabad, Telangana 500014
shravanreddybejjenki@gmail.com[1]hodece@mriet.ac.in[2]rajakishore.r@gmail.com[3]

Abstract---*Content-aware image retrieval is a very important topic nowadays, when the amount of digital image data is highly increasing. Existing PCA (principle component analysis) based image retrieval systems perform at a reduced level on real life images, where background data may distort image descriptors and retrieval results. To avoid this, a preprocessing step is introduced in this paper to distinguish between foreground and background, using integrated saliency detection. To build the descriptor only on the most relevant pixels, orientation feature is extracted at salient Modified Harris for Edges and Corners (MHEC) key points using an improved edge map, resulting in a Salient Orientation Histogram (SOH). The proposed CBIR system is also augmented with a segmentation step for object detection. The method is tested on the CORAL database, containing random internet images. Image retrieval and object detection both give promising results compared to other state-of-the-art methods. In this study, PCA based image retrieval system and proposed CBIR are compared for performance analysis. The proposed method yields good performance results.*

Index Terms— Direction selectivity, modified Harris for edges and corners, saliency detection, SBIR

1. INTRODUCTION

Content Based Image Retrieval (CBIR) is a powerful tool. It uses the visual cues to search images databases and retrieve the required images. It uses several approaches and techniques for this purpose. The visual contents of images, such as color, texture, shape and region, are extensively explored for indexing and representation of the image contents. These low level features of an image are directly related to the contents of the image. These image contents could be extracted from image and could be used for measuring the similarity amid the queried image and images in the database using different statistical methods. In content-based retrieval systems different features of an image query are exploited to search for analogous images features in the database.

Content-aware image retrieval is a very important topic nowadays with the constantly increasing amount of digital image data. Outline sketches have recently been shown to be more comfortable for retrieval than a complete image, as sketch based image retrieval (SBIR) expects simpler descriptors resulting in faster comparison and retrieval. Descriptors can be grouped into global and local types. While the former includes information of the whole image, the latter concentrates only on a small image part. Recently published SBIR systems employs local features, as global ones are not handling affine variations well, and the fact that fine details of the drawing are often missing.

Existing SBIR systems are mainly tested on image databases without significant background information. However, randomly selected internet images often contain a lot of background data with varying texture and color, which can influence the image descriptors and make the comparisons more challenging. To avoid this, a preprocessing step can help to distinguish between foreground and background, which increases the importance of saliency detection. However, the dimension of a salient area description can still be very high, thus further reduction is needed. Interest point detectors, like Harris emphasize relevant structures in the image.

Thus, if the local descriptors are calculated at interest point locations, the extracted salient region information can be reduced while retaining their relevance. Modified Harris for Edges and Corners (MHEC) was proposed earlier by the author for efficient image segmentation, and the method's strong ability for object detection was also shown previously, supporting its capability of holding efficient structure and content information for image comparisons and retrieval. Orientation as a descriptor has already been introduced in earlier SBIR systems; moreover many improvements of the Histogram of Oriented Gradients (HoG) were published over the past years.

The original HoG calculated the histogram for the whole image. Improved adaptations of HoG for SBIR systems are mostly using canny edge maps with orientation histograms calculated on pixels of the Canny edge map or randomized pixels. Following this technique, the background texture maycreate false edges in the canny edge map and the keypoint selection could include background hits. Both of them may cause the distortion of the orientation histogram and reduced retrieval accuracy.

2. RELATED WORK

Research on CBIR could be bifurcated into two groups on the basis of the features used to retrieve the required image. Early approaches used a single feature out of the available features namely shape, texture, color and region for retrieval of the required image. Results of single feature based retrieval systems were not satisfactory because generally image contains several visual features. The current approaches use different combination of the visual features to retrieve the required image.

The shape descriptor also provides dominant information in image retrieval because shape is the only source through which humans can recognize objects. The shape feature can be retrieved by two methods boundary based shape feature extraction and region based shape extraction. The boundary based technique is based on outer boundary while the region based technique is depending on the whole region.

An efficient CBIR system with better performance is presented by using the wavelets decomposition of image; they have generated the composite sub-band gradient and the energy distribution pattern string from the sub images of are generated by means of wavelet decomposition to the input image. For filtering out the undesired images a technique based on energy distribution pattern strings fuzzy matching is used. The resultant images are compared with query image after filtering. The system is tested on the database of 2400 images.

3. LITERATURE REVIEW

A typical image retrieval system includes three major components: i) feature extraction (usually in conjunction with feature selection), ii) high dimensional indexing and iii) system design [3]. An image can be represented as a set of low- level visual features such as color, texture and shape features. While several image retrieval systems rely on only one feature for the extraction of relevant images, it has been shown that an appropriate combination of relevant features can yield better retrieval performance [4]. The process of determining the combination of features that is most representative of a particular query image is called feature selection. Works has been done on color and texture feature extraction algorithms. Feature selection algorithm based on fuzzy approach and relevance feedback has been given.

3.1Color Feature Extraction Color features include the conventional color histogram (CCH), the fuzzy color histogram (FCH), the color correlogram (CC) and a more recent color- shape- based feature. The extraction of the color- based features follows a similar progression in each of the four methods: i) Selection of the color space, ii) quantization of the color space, iii) extraction of the color feature, iv) derivation of an appropriate distance function.

3.2Texture Feature Extraction Texture feature extraction methods include the steerable pyramid, the contourlet transform, the Gabor wavelet transform and the complex directional filter bank (CDFB).

3.3Shape Feature Extraction In image retrieval, as per applications, shape representation are required to be either invariant to translation, rotation, and scaling or not. Hence, two categories of shape representations can be distinguished, boundary and region based. The first utilizes only the outer boundary of the shape while the other access the entire shape region [10]. The most successful representatives for these two categories are Fourier descriptor and moment invariants. The main idea of a Fourier descriptor is to use the Fourier transformed boundary as the shape feature. Some early work can be found in [11, 12]. Rui et al. proposed a modified Fourier descriptor which is both robust to noise and invariant to geometric transformations [40]. The main idea of moment invariants is to use region-based moments which are invariant to transformations, as the shape feature.

3.4Fuzzy Feature Selection With Relevance Feedback The goal of feature selection is to find the optimal feature subspace where the 'relevant' and 'irrelevant' feature sets are best separated. In an attempt to bridge the gap between high- level user semantics and low- level visual features, Jiang et al. proposed in [26] an online feature selection algorithm in the relevance feedback learning process. The online feature selection algorithm is implemented in a boosting manner by combining incrementally learned classifiers over the selected features into a strong ensemble classifier. The learning phase involves acquiring feedback from users that are asked to label the initially returned images as 'relevant' or 'irrelevant'.

3.5Content Based Image Retrieval By Multi Features The earliest work on Content Based Image Retrieval was done by Ning-San Chang and King-Sun Fu in their paper Query-by-Pictorial-Example. They introduced Query-by-Pictorial-Example as a relational query language for manipulating queries regarding pictorial relations as well as conventional relations.

T. Joseph, A.F. Cardenas presented a corresponding high-level query language, PICQUERY. Eden and Unse described an approach where local statistics (texture energy measures) are estimated at the output of an equivalent filter bank by means of a nonlinear transformation (absolute value) followed by an iterative Gaussian smoothing algorithm.

4. PROPOSED METHOD

Image Retrieval and Segmentation Here for image retrieval we are using Scale Invariant Feature Transform.

4.1. Scale-invariant feature transform (SIFT):

SIFT is an algorithm in computer vision to detect and describe local features in images. For any object in animage, interesting points on the object can be extracted to provide a "feature description" of the object. This description, extracted from a training image, can then be used to identify the object when attempting to locate the object in a test image containing many other objects. To perform reliable recognition, it is important that the features extracted from the training image be detectable even under changes in image scale, noise and illumination. Such points usually lie on high-contrast regions of the image, such as object edges. Another important characteristic of these features is that the relative positions between them in the original scene shouldn't change from one image to another.

A. Scale-space extrema detection:

We begin by detecting points of interest, which are termed keypoints in the SIFT framework. The image is convolved with Gaussian filters at different scales, and then the difference of successive Gaussian-blurred images are taken. Key points are then taken as maxima/minima of the Difference of Gaussian(DoG) that occur at multiple scales. Specifically, a DoG image is given by where is the convolution of the original image with the Gaussian blur at scale i.e. Hence a DoG image between scales and is just the difference of the Gaussian-blurred images at scales and For scale space extrema detection in the SIFT algorithm, the image is first convolved with Gaussian-blurs at different scales. The convolved images are grouped by octave (an octave corresponds to doubling the value of), and the value of is selected so that we obtain a fixed number of convolved images per octave. Then the Difference-of-Gaussian images are taken from adjacent Gaussian-blurred images per octave.

B.Keypoint localization:

Scale-space extrema detection produces too many keypoint candidates, some of which are unstable. The next step in the algorithm is to perform a detailed fit to the nearby data for accurate location, scale, and ratio of principal curvatures. This information allows points to be rejected that have low contrast (and are therefore sensitive to noise) or are poorly localized along an edge. First, for each candidate keypoint, interpolation of nearby data is used to accurately determine its position. The initial approach was to just locate each keypoint at the location and scale of the candidate keypoint. The new approach calculates the interpolated location of the extremum, which substantially improves matching and stability.

C. Get rid of Low Contrast points:

Key points generated in the previous step produce a lot of key points. Some of them lie along an edge, or they don't have enough contrast. In both cases, they are not useful as features. So we get rid of them. The approach is similar to the one used in the Harris Corner Detector for removing edge features. This is simple, If the magnitude of the intensity (i.e., without sign) at the current pixel in the DoG image (that is

International Journal of Research
Available at https://edupediapublications.org/journals
e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 04 Issue-17
December 2017

being checked for minima/maxima) is less than a certain value, it is rejected. Because we have subpixel keypoints (we used the Taylor expansion to refine keypoints), we again need to use the taylor expansion to get the intensity value at subpixel locations. If it's magnitude is less than a certain value, we reject the keypoint.

D. Salient Oriented Histogram:

After step 3, we have legitimate key points. They've been tested to be stable. We already know the scale at which the keypoint was detected (it's the same as the scale of the blurred image). So we have scale invariance. The next thing is to assign an orientation to each keypoint. This orientation provides rotation invariance. The idea is to collect gradient directions and magnitudes around each keypoint. Then we figure out the most prominent orientation(s) in that region. And we assign this orientation(s) to the keypoint. Any later calculations are done relative to this orientation. This ensures rotation invariance.

The size of the "orientation collection region" around the keypoint depends onit's scale. The bigger the scale, the bigger the collection region.

First, the Gaussian-smoothed image$L(x,y,\sigma)$ at the keypoint's scale $\sigma$is taken so that all computations are performed in a scale-invariant manner. For an image sample $L(x,y)$ at scale σ, the gradient magnitude, $m(x,y)$, and orientation, $\theta(x,y)$, are precomputed using pixel differences: Gradient magnitudes and orientations are calculated using these formulae:

$$m(x,y) = \sqrt{L(x+1,y) - L(x-1,y)^2 + (L(x,y+1) - L(x,y-1))^2}$$

$$\theta(x,y) = \tan^{-1}\left(\frac{L(x,y+1) - L(x,y-1)}{L(x+1,y) - L(x-1,y)}\right) \qquad (2)$$

The magnitude and orientation is calculated for all pixels around the keypoint. Then, A histogram is created for this.

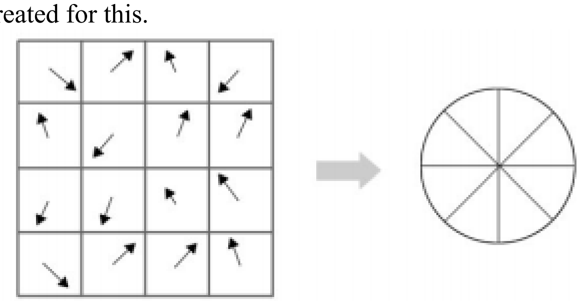


Fig.1. Orientation Assignment

In this histogram, the 360 degrees of orientation are broken into 36 bins (each 10 degrees). Lets say the gradient direction at a certain point (in the "orientation collection region") is 18.759 degrees, then it will go into the 10-19 degree bin. And the "amount" that is added to the bin is proportional to the magnitude of gradient at that point. Once you've done this for all pixels around the keypoint, the histogram will have a peak at some point. Above, you see the histogram peaks at 20-29 degrees.

So, thekeypoint is assigned orientation 3 (the third bin). Also, any peaks above 80% of the highest peak are converted into a new keypoint. This new keypoint has the same location and scale as the original. But it's orientation is equal to the other peak. So, orientation can split up one keypoint into multiple keypoints.

E.Keypoint descriptor:

Previous steps found keypoint locations at particular scales and assigned orientations to them. This ensured invariance to image location, scale and rotation. Now we want to compute a descriptor vector for each keypoint such that the descriptor is highly distinctive and partially invariant to the remaining variations such as illumination, 3D viewpoint, etc. This step is performed on the image closest in scale to the keypoint's scale.

(

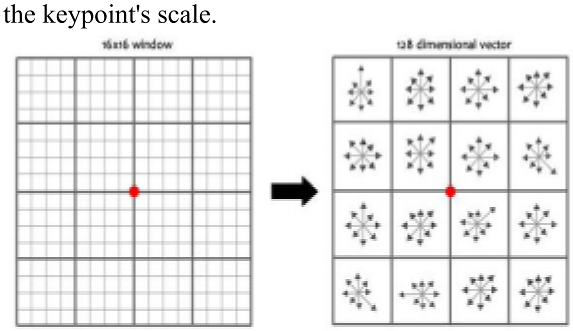


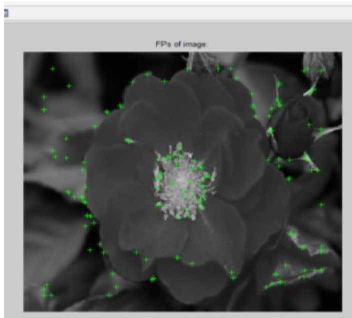
Fig.2: locating the descriptor points.

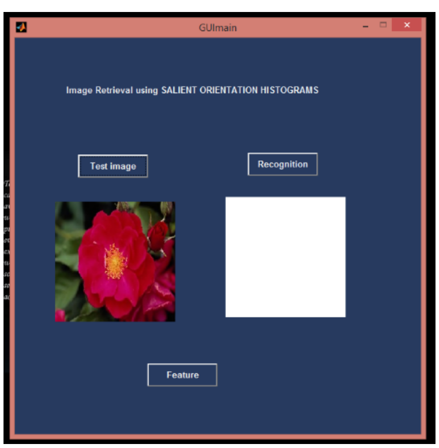
First a set of orientation histograms is created on 4x4 pixel neighborhoods with 8 bins each. These histograms are computed from magnitude and orientation values of samples in a 16 x 16 region around the keypoint such that each histogram contains samples from a 4 x 4 subregion of the original neighborhood region. The magnitudes are further weighted by a Gaussian function with equal to one half the width of the descriptor window. The descriptor then becomes a vector of all the values of these histograms. Since there are 4 x 4 = 16 histograms each with 8 bins the vector has 128

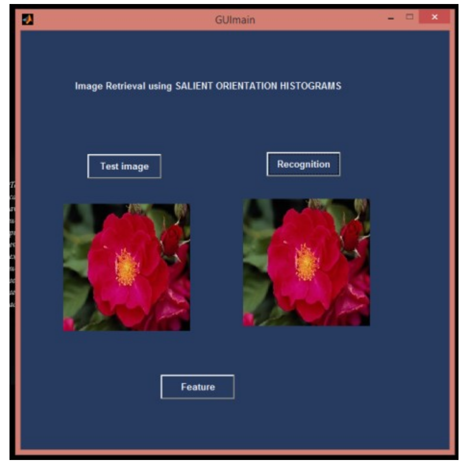
e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 04 Issue-17
December 2017

elements. This vector is then normalized to unit length in order to enhance invariance to affine changes in illumination. To reduce the effects of non-linear illumination a threshold of 0.2 is applied and the vector is again normalized.

5. SIMULATION RESULTS







6. CONCLUSION

In this paper, a novel SBIR system is introduced, using a salient keypoint based orientation histogram (SOH). The proposed method first extracts the salient image region based on texture distinctiveness, followed by a Modified Harris for Edges and Corners (MHEC) interest point detection. This way the most relevant pixels of the image are selected to build an orientation histogram on an improved edge map, instead of applying Canny edge map like earlier SBIR systems. The edge map is also adapted for segmentation. Overall, the proposed descriptor achieves high performance on the THUR15000 dataset, and it also provides an efficient object detection method. Future work will investigate the improved integration of saliency in SBIR systems.

REFERENCES

[1] AbdolahChalechale, GolshahNaghdy, and Alfred Mertins, "Sketch-based image matching using angular partitioning," Systems, Man and Cybernetics, Part A: Systems and Humans, IEEE Transactions on, vol. 35, no. 1, pp. 28–41, 2005.

[2] David G. Lowe, "Object recognition from local scaleinvariant features," in Proc. of International Conf. on Computer Vision, 1999, pp. 1150–1157.

[3] Mathias Eitz, Kristian Hildebrand, TamyBoubekeur, and Marc Alexa, "Sketch-based image retrieval: Benchmark and bag-of-features descriptors," Visualization and

Computer Graphics, IEEE Transactions on, vol. 17, no. 11, pp. 1624–1636, 2011.
[4] Ming Cheng, Niloy J Mitra, Xumin Huang, Philip HS Torr, and Song Hu, "Global contrast based salient region detection," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 37, no. 3, pp. 569– 582, 2015.
[5] Christian Scharfenberger, Alexander Wong, Khalil Fergani, John S Zelek, David Clausi, et al., "Statistical textural distinctiveness for salient region detection in natural images," in Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on, 2013, pp. 979– 986.
[6] EsaRahtu, JuhoKannala, MikkoSalo, and Janne Heikkila, "Segmenting salient objects from images and ¨ videos," in European Conference on Computer Vision, pp. 366–379. Springer, 2010.
[7] Radhakrishna Achanta, Sheila Hemami, Francisco Estrada, and Sabine Ssstrunk, "Frequency-tuned Salient Region Detection," in IEEE International Conference on Computer Vision and Pattern Recognition (CVPR 2009), 2009, pp. 1597 – 1604.
[8] C. Harris and M. Stephens, "A combined corner and edge detector," in Proc. of the 4th Alvey Vision Conf., 1988, pp. 147–151.
[9] A. Kovacs and T. Szir ´ anyi, "Improved Harris feature ´ point set for orientation sensitive urban area detection in aerial images," IEEE Geoscience and Remote Sensing Letters (in press), vol. 10, no. 4, pp. 796–800, 2013.
[10] A. Manno-Kovacs, "Direction selective vector field convolution for contour detection," in IEEE International Conference on Image Processing (ICIP), 2014, pp. 4722–4726, "Top10%" award.
[11] Navneet Dalal and Bill Triggs, "Histograms of oriented gradients for human detection," in Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, 2005, vol. 1, pp. 886–893.
[12] Rui Hu, Mark Barnard, and John Collomosse, "Gradient field descriptor for sketch based retrieval and localization," in Image Processing (ICIP), 2010 17th IEEE International Conference on. IEEE, 2010, pp. 1025–1028.
[13] Ming-Ming Cheng, NiloyJ. Mitra, Xiaolei Huang, and Shi-Min Hu, "Salientshape: group saliency in image collections," The Visual Computer, vol. 30, no. 4, pp. 443–453, 2014.
[14] N. Otsu, "A threshold selection method from gray-level histograms," IEEE Transactions on Systems, Man and Cybernetics, vol. 9, no. 1, pp. 62–66, 1979.
[15] Andrea Kovacs and Tamas Sziranyi, "Harris function based active contour external force for image segmentation," Pattern Recognition Letters, vol. 33, no. 9, pp. 1180–1187, 2012.
[16] Sanjiv Kumar and Martial Hebert, "Man-made structure detection in natural images using a causal multiscale random field," in Proc. IEEE Conf. Comput. Vision Pattern Recogn., 2003, pp. 119–126.
[17] Ofir Pele and Michael Werman, "The quadratic-chi histogram distance family," in Proceedings of the 11th European Conference on Computer Vision: Part II. 2010, pp. 749–762, Springer-Verlag.