Volume 2, Issue 4, April 2012 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

Radial Basis Function used in CBIR for SIFT Features

Shraddha Kumar*
Department of IT
JIMS, Sec-5, Rohini, GGSIPU
shraddha20.4@gmail.com

Shweta Singhal
Department of IT
JIMS, Sec-5, Rohini, GGSIPU

Ankita Sharma
Department of IT
JIMS, Sec-5, Rohini, GGSIPU

ABSTRACT— ***The traditional information retrieval techniques does not meet the user's demand, so there is need to develop an efficient system for content based image retrieval (CBIR). In this paper method is proposed for an efficient CBIR system using Radial basis function (RBF) which can give better results than other existing CBIR. The proposed system use the Scale Invariant Feature Transform (SIFT) to extract the feature from the database. Illumination changes does not affect the system performance due to these SIFT features. These features generate keypoints. The distances between feature vectors of database and the query image are computed using the classifier RBF.***

Keywords— **Image retrieval, Object Recognition, CBIR, SIFT, RBF.**

I. Introduction

CBIR systems describe each image (either the query or the ones in the database) by a set of features that are automatically extracted. Then, the feature vectors are fed into a classifier. The processes of image feature selection and extraction uses descriptors, such as Scale Invariant Feature Transform (SIFT). SIFT (Scale Invariant Feature Transform) features are widely used in object recognition. These features are invariant to changes in scale, 2D translation and rotation transformations [1]. To a limited extent they are also robust to 3D projection transformations. SIFT Features however, are of very high dimension and large number of SIFT features are generated from an image. The large computational effort associated with matching all the SIFT features for recognition tasks, limits its application to object recognition problems. Image matching is a fundamental aspect of many problems in computer vision, including object or scene recognition, solving for 3D structure from multiple images, stereo correspondence, and motion tracking. This paper describes image features that have many properties that make them suitable for matching differing images of an object or scene. The features are invariant to image scaling and rotation, and partially invariant to change in illumination and 3D camera viewpoint [3, 4]. They are well localized in both the spatial and frequency domains, reducing the probability of disruption by occlusion, clutter, or noise. Large numbers of features can be extracted from typical images with efficient algorithms. In addition, the features are highly distinctive, which allows a single feature to be correctly matched with high probability against a large database of features, providing a basis for object and scene recognition. Matching of the features is done by the Radial basis function (RBF) [6]. It is the approach of the neural network which is used for the matching of input from the output.

II. Proposed Framework

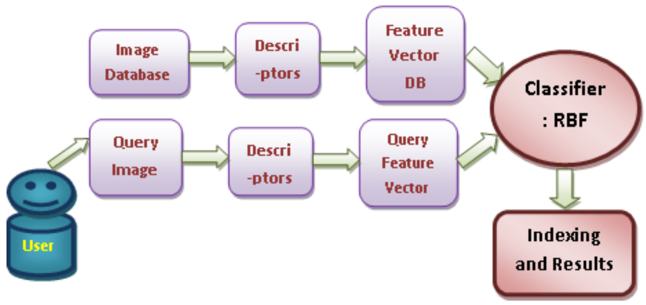

Fig. 1 Proposed Framework

In the proposed framework (fig. 1) we are having image database and the query image. We are using the SIFT features to get the features of the image. We get descriptors and the feature vector by using SIFT algorithm. These keypoints are matched by using a classifier i.e. RBF function which is used for the similarity matching from the database. After matching the image is retrieved from the large database.

Feature Extraction

Following are the major stages of computation used to generate the set of image features:

1. **Scale-space extrema detection:** The first stage of computation searches over all scales and image locations. It

is implemented efficiently by using a difference-of-Gaussian function to identify potential interest points that are invariant to scale and orientation [1].

2. **Keypoint localization:** At each candidate location, a detailed model is fit to determine location and scale. Key points are selected based on measures of their stability [1].

3. **Orientation assignment:** One or more orientations are assigned to each key point location based on local image gradient directions. All future operations are performed on image data that has been transformed relative to the assigned orientation, scale, and location for each feature, thereby providing invariance to these transformations [1].

4. **Keypoint descriptor:** The local image gradients are measured at the selected scale in the region around each keypoint. These are transformed into a representation that allows for significant levels of local shape distortion and change in illumination [1].

This approach has been named the Scale Invariant Feature Transform (SIFT), as it transforms image data into scale-invariant coordinates relative to local features. For image matching and recognition, SIFT features are first extracted from a set of reference images and stored in a database. A new image is matched by individually comparing each feature from the new image to this previous database and finding candidate matching features based on Euclidean distance of their feature vectors. Fast nearest-neighbor algorithms are also used for matching the SIFT features [5]. The keypoint descriptors are highly distinctive, which allows a single feature to find its correct match with good probability in a large database of features. However, in a cluttered image, many features from the background will not have any correct match in the database, giving rise to many false matches in addition to the correct ones. The correct matches can be filtered from the full set of matches by identifying subsets of keypoints that agree on the object and its location, scale, and orientation in the new image. The probability that several features will agree on these parameters by chance is much lower than the probability that any individual feature match will be in error. The determination of these consistent clusters can be performed rapidly by using an efficient hash table implementation of the generalized Hough transform [2].
The SIFT feature algorithm is based upon finding locations within the scale space of an image which can be reliably extracted. The first stage finds scale-space extreme located in D(x, y, θ), the Difference of Gaussians (DOG) function, which can be computed from the difference of two nearby scaled images separated by a multiplicative factor k:

$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y) = L(x, y, k\sigma) - L(x, y, \sigma)$$

where L(x, y, σ) is the scale space of an image, built by convolving the image I(x, y) with the Gaussian kernel G(x, y, σ). Points in the DOG function which are local extreme in their own scale and one scale above and below are extracted as keypoints. Generation of extreme in this stage is dependent on the frequency of sampling in the scale space k and the initial smoothing σ. The keypoints are then filtered for more stable matches, and more accurately localized to scale and sub-pixel image location using methods.

Before a descriptor for the keypoint is constructed, the keypoint is assigned an orientation to make the descriptor invariant to rotation. This keypoint orientation is calculated from an orientation histogram of local gradients from the closest smoothed image L(x, y, σ). For each image sample L(x, y) at this scale, the gradient magnitude m(x, y) and orientation θ(x, y) is computed using pixel differences:

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

$$\theta(x, y) = \tan^{-1}((L(x, y+1) - L(x, y-1)) / (L(x+1, y) - L(x-1, y)))$$

The orientation histogram has 36 bins covering the 360 degree range of orientations. Each point is added to the histogram weighted by the gradient magnitude, m(x, y), and by a circular Gaussian with σ variance that is 1.5 times the scale of the keypoint. Additional keypoints are generated for keypoint locations with multiple dominant peaks whose magnitude is within 80% of each other [1].
The dominant peaks in the histogram are interpolated with their neighbors for a more accurate orientation assignment.

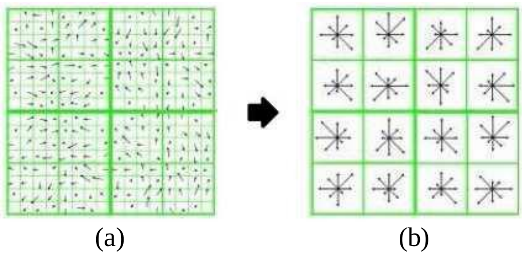

Fig. 2 A sample SIFT descriptor computation. (a) The gradients of an image patch around a keypoint. These gradients are then accumulated over *4 x 4* sub-regions, as shown on the (b).The length of the arrow corresponds to the sum of the gradient magnitudes in that direction.

Fig.2 represents a sample SIFT descriptor of the shown image. The local gradient data from the closest smoothed image L(x, y, σ) is also used to create the keypoint descriptor. This gradient information is first rotated to align it with the assigned orientation of the keypoint and then weighted by a Gaussian with σ variance that is 1.5 times the scale of the keypoint. The weighted data is used to create a nominated number of histograms over a set window around the keypoint use 16 orientation histograms aligned in a 4x4 grid. Each histogram has 8 orientation bins each created over a support window of 4x4 pixels. The resulting feature vectors are 128 elements with a total support window of 16x16 scaled pixels.

III. SIMILARITY MATCHING

In the proposed method we are using the Radial Basis Function (RBF) for matching the images from the database. RBF is embedded in a two layer neural network, where each hidden unit implements a radial activated function. The output unit implements a weighted sum of hidden unit outputs [7]. The input into an RBF network is nonlinear while the output is linear. Due to their nonlinear approximation properties, RBF network are able to model complex mappings, which perceptron neural networks can only model by means of multiple intermediary layers [8].
In order to use a Radial Basis Function Network we need to specify the hidden unit activation function the number of processing units, a criterion for modeling a given task and a training algorithm for finding the parameters of the network. Finding the RBF weights is called network training [6].
A radial basis function (RBF) is a real-valued function whose value depends only on the distance from the origin, so that $\Phi(X) = \Phi(\|X\|)$; or alternatively on the distance from some other point *c*, called a *center*, so that $\Phi(X, C) = \Phi(\|X-C\|)$.Any function φ that satisfies the property $\Phi(X) = \Phi(\|X\|)$ is a radial function. The norm is usually Euclidean distance, although other distance functions are also possible [6].

Fig. 3 Matching Image by Keypoints

In Fig.3 there is shown the matching procedure of the keypoints retrieved by the SIFT features. The matching is done by the RBF function. The keypoints generated by the SIFT are trained by the supervised training used in neural network. Radial basis function methods have their origins in techniques for performing exact interpolation of a set of data points in a multi- dimensional space. The exact interpolation problem requires every input vector, and forms a convenient starting point.

IV. CONCLUSIONS

The SIFT keypoints described in this paper are particularly useful due to their distinctiveness, which enables the correct match for a keypoint to be selected from a large database of other keypoints. Large numbers of keypoints can be extracted from typical images, which leads to robustness in extracting small objects among clutter. The fact that keypoints are detected over a complete range of scales means that small local features are available for matching small and highly occluded objects, while large keypoints perform well for images subject to noise and blur. This paper presents the keypoints recognition by using the Radial Basis Function. The keypoints generated are large in numbers so this matching process may take a large time to retrieve the image from the database. To reduce the time we should select some of the keypoints and match them. These is work may be done in future.

REFERENCES

[1]. D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, 60, 2 (2004), pp. 91-110.
[2]. A. Suga, K. Fukuda, T. Takiguchi Y. Ariki, "Object recognition and segmentation using SIFT and Graph Cuts", *ICPR*, 2008, pp. 1-4.
[3]. Brown, M. and Lowe, D.G. 2002. Invariant features from interest point groups. In *British Machine Vision Conference*, Cardiff, Wales, pp. 656-665.
[4]. M. Brown and D.G. Lowe, "Invariant features from interest point groups," In *British Machine Vision Conference*, BMVC 2002, Cardiff, Wales (September 2002), pp. 656-665.
[5]. D.G.Lowe, "Object recognition from local scale invariant features," *International Conference on Computer*.
[6]. A.G. Bors, "Introduction to Radial Basis Function (RBF) Networks." Department of computer science University of York.
[7]. S. Haykin. "*Neural Networks: a comprehensive foundation*". Tom Robins, 2nd edition, 1999.
[8]. Y. Zhang, X. Rong Li, Z. Zhu, and H. Zhang, "A new clustering and training method for radial basis function networks," in *Proc. of the Intern. Conf. on Neural Networks - ICNN*, pp. 311–316, 1996.

