

# Region based Image Similarity using Fuzzy based SIFT Matching

Shashikanth CC  
ADE, Bangalore

Parag Kulkarni  
EkLAT Research, Pune

## ABSTRACT

This paper proposes a region based approach using Scale Invariant Feature Transform (SIFT) and fuzzy logic for computing region based image similarity. When SIFT algorithm is used for matching images, a meaningful or semantic match according to human perception is not obtained. Hence to refine its output, a region based approach has been proposed. When a test image is compared with reference image, SIFT descriptors are computed and the images are segmented into regions and labeled. SIFT similarity measure along with the region information is given as input to fuzzy logic to determine region based similarity measure. Experiments are done using real world optical images, Caltech image datasets and a few satellite images. The proposed approach is found to be efficient for optical images, Caltech image datasets and have good performance for satellite images.

## KEYWORDS

SIFT descriptors, Fuzzy logic, Region based image similarity

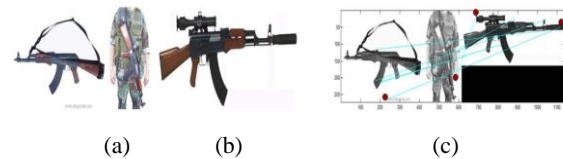
## 1. INTRODUCTION

Due to the digitization of data and advances in technology, object recognition and content based image retrieval systems became important research areas. Image similarity had a key role in both these systems. Features extracted from images can be classified into two broad categories: Local and Global features. Global features varies from simple statistical measures such as color, texture, shape [1] to more sophisticated dimensionality reduction techniques [2]. Global features are non robust to occlusion and lack of invariance, whereas local features are robust to occlusion and clutter. Some of the local feature descriptors [3] are Harris corner based detector [4], Hessian matrix, Scale Invariant Feature Transform (SIFT) [5], Gradient Location Orientation Histogram, edge based region detector [6] etc. Design of a suitable similarity measure depends on many factors such as agreement with semantics, robust to noise etc [7].

SIFT descriptors are invariant to scaling, translation, rotation, and partially invariant to illumination changes and affine or 3D projection. In SIFT based matching, the matched keypoints will not always be correctly paired to corresponding regions of the test and reference image. In this paper, an efficient means for correctly matching the keypoints based on its region, has been proposed. The fuzzy logic for region based matching has been proposed. Experiments were done using real-world optical images, Caltech image datasets and satellite images. The proposed algorithm is found to perform better than SIFT algorithm in discriminating similar and dissimilar images. Section 2 explains the proposed algorithm in detail. Experimental results are discussed in section 3 and section 4 concludes the paper.

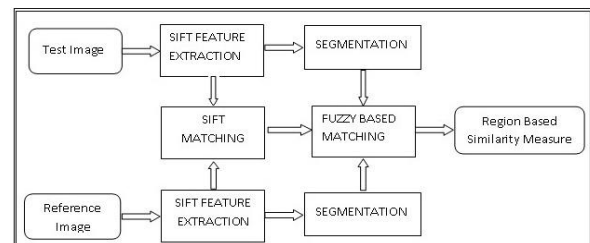
## 2. PROPOSED REGION BASED ALGORITHM

The objective of this proposed system is to find similarity between two images according to human perception. The proposed method is a region based approach which refines the output of SIFT matching between images. SIFT [5] based analysis involves detecting salient locations in an image and extracting descriptors that are distinctive and invariant to changes in viewpoint, illumination, scale, noise and local geometric distortion, for performing reliable recognition. The motivation for this region based algorithm is, when SIFT based matching is applied, the matched keypoints will not always correspond to same regions of the test and reference image. An example is shown in Figure 1, where some of the keypoints in test image are mismatched with the keypoints of the object 'AK47' in the reference image. This region based approach is used to refine the output obtained from SIFT matching.



**Figure 1 An example which illustrates mismatched points between different objects using SIFT Algorithm. (a) Test image; (b) Reference image; (c) SIFT matching. The mismatched points between different objects are marked with circles.**

The overall architecture is given in Figure 2. SIFT descriptors are computed from test and reference images and segmented based on the location of keypoints. The segmented regions are labeled and each keypoint is assigned its corresponding region number. SIFT matching is performed to find matched keypoints and similarity measure between them. For each matched keypoint pair, SIFT similarity measure and belongingness of keypoints to its regions are given as input to fuzzy system to determine region based similarity measure.

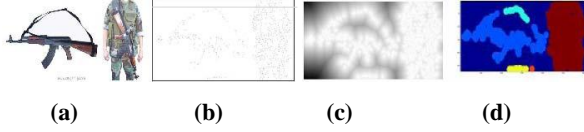


**Figure 2: Overall Architecture of the proposed system**

### 2.1 Region based Segmentation

Segmentation is performed using the distance transform of the binary image, which depicts the location of keypoints. Figure

3 (b) shows the location of key points (as black pixels) of the input image in Figure 3 (a). Distance transform is applied on this binary image.



**Figure 3: Segmentation Steps: (a) Input image; (b) Location of keypoints (marked as black pixels) (c) Distance transform obtained based on location of SIFT descriptors ; (d) Labeled output after applying threshold on (c).**

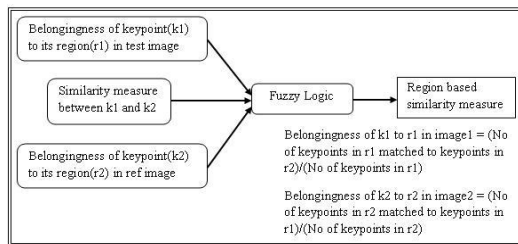
Figure 3 (c) shows the distance transformed image of the Figure 3 (b). After performing distance transform, a threshold is chosen to segment the image into different regions . The thresholded image is labeled using connected component labeling. The labeled output of Figure 3 (c) is shown in Figure 3 (d). The input image is now segmented into different regions / objects.

## 2.2 Fuzzy based Matching

The fuzzy logic to has been designed determine the degree of similarity between test and reference images. The inputs to the fuzzy logic for every pair of matched keypoints (k1,k2) obtained using SIFT matching are as follows:

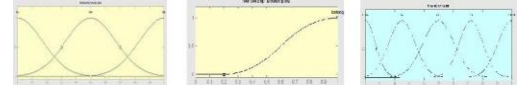
- 1 Similarity measure obtained using standard SIFT matching
- 2 Belongingness of keypoint (k1) to its region (r1) in test image
- 3 Belongingness of keypoint (k2) to its region (r2) in reference image

The belongingness of a keypoint in r1 is given by the ratio between the number of keypoints in r1 matched to number of keypoints in r2 and the number of in r1. Similarly belongingness of keypoint (k2) to r2 is also defined as vice versa. The block diagram of fuzzy based matching is shown in Figure 4. The details of fuzzy logic are described below. The membership functions for input variable 'SIFT similarity measure' is defined by Gaussian functions for 'low',



**Figure 4: Block diagram of fuzzy based matching.**

'medium' and 'high' (see Figure 5 (a)). The input variable 'belongingness' is defined by sigmoid function, shown in Figure 5 (b). Output variable is defined by Gaussian functions for 'very low', 'low', 'med', 'high' and 'very high' (see Figure 5 (c))



**Figure 5: Fuzzy membership functions for (a) SIFT similarity measure (b) Belongingness to the region (c) Output Region based similarity measure.**

The following are the fuzzy rules used for the proposed system to determine the region based similarity measure. The defuzzification method used for the output is centroid method.

- 1 If (SIFT measure is low) then (similarity measure is very low)
- 2 If (SIFT measure is medium) and (r1 is belong) and ( r2 is belong ) then (similarity measure is high)
- 3 If (SIFT measure is high) then (similarity measure is very high)
- 4 If (SIFT measure is medium) and (r1 is not belong) and ( r2 is not belong ) then (similarity measure is low)
- 5 If (SIFT measure is medium) and (r1 is not belong) and ( r2 is belong ) then (similarity measure is medium)
- 6 If (SIFT measure is medium) and (r1 is belong) and ( r2 is not belong ) then (similarity measure is medium)

**Table 1. Region based similarity measure for images undergone rotation.**

Test image	Reference image: test image rotated by angle (in degrees)				
	0	45	90	135	180
Region based similarity measure	1	0.7821	0.9707	0.7758	0.9574

For each matched keypoints of test image using SIFT, a region based similarity measure is obtained using fuzzy logic (FSM) discussed in previous section. The keypoints are said to be semantically matched if they lie above a threshold and unmatched otherwise. The final similarity measure (RSM) is obtained using the following equation.

$$RSM = \frac{k}{n} * 0.25 + \frac{1}{k} \sum_{i=1}^k FSM(k_i) * 0.75 \quad \text{---- (1)}$$

where k is the number of matched keypoints; n is the total number of keypoints in test image and FSM is the similarity measure obtained from fuzzy logic.

## 3. EXPERIMENTAL RESULTS AND DISCUSSION

The experiment has been carried out with three different types of datasets; Dataset 1: Caltech database [8]; Dataset 2: Real

world images [9]; Dataset 3: Image shots from SONY Camcorder; and a few aerial images. This approach is tested with similar and dissimilar images to check its efficiency in discriminating images and also compared with SIFT approach. And the system is also experimented by varying its threshold to find the optimum threshold for different datasets.









### 3.1 Testing the performance

The normalized similarity measure (computed using the Equation 1) gives the degree of similarity between test and reference image. Higher the value of similarity measure, the test and reference images are more similar.


The proposed method is invariant to rotation, scaling and illumination changes. In Table 1, reference images are generated by rotating to angles of  $45^{\circ}$ ,  $90^{\circ}$ ,  $135^{\circ}$  and  $180^{\circ}$  using bilinear interpolation. The similarity measure lies between 0.77 to 1. This proposed method is tested for invariance to scaling/illumination changes and perspective projections (see Table 2). The similarity measure lies in the range of 0.6 to 0.7.

Another set of experiments was done using images which contain different number of objects. In first row of table 3, the reference images are ordered in their increasing order of similarity from left to right. The similarity measure increases from left to right except for the third reference image. This is due to the reason that the object 'pen' in the third reference image is exactly same as the test image whereas the object 'pen' in fourth reference image is different from the test image. This shows that the algorithm is more tuned towards similarity between objects rather than the number of objects in the images. This approach has been tested on satellite images (second row of table 3). This time series of photographs [10] of the Mediterranean coast near the Suez Canal shows ten years of changes associated with two of Egypt's "Mega" Development Projects. The proposed approach has detected similarity between images with value of 0.59. The above mentioned experiments show that similar images have high similarity measure.

**Table 2. Region based similarity measure for images with illumination changes.**

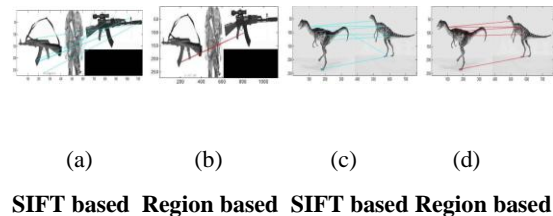
Test image	Reference images and its region based similarity measure	
		0.6814
		0.6315
		0.6195
		0.5977

**Table 3. Region based similarity measure between test image and reference images with different no. of objects. Satellite images are taken from NASA website [10].**

Test image	Reference images and its region based similarity measure				
					
	0.5732	0.5972	0.6360	0.6201	
					
	0.5828	0.5866	0.5980	0.5956	

### 3.2 Comparison between Region based and SIFT matching

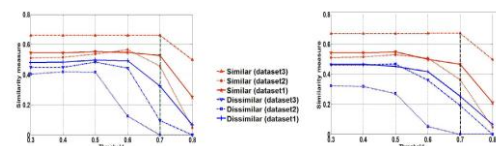
In this section, it has been illustrated that, the performance of this proposed region based approach is better than the standard SIFT algorithm. Figure 6 shows two different examples comparing region based with SIFT based approach. The mismatched keypoints using standard SIFT based approach is removed in the region based approach, while preserving the correctly matched keypoints.



**Figure 6: Examples illustrating the better performance of the proposed region based approach than standard SIFT matching.**

### 3.3 Optimal threshold determination

To determine the optimal threshold for region based matching, for each dataset, a test image is compared with 50 other reference images which belongs to the same class as test image (similar images). Similarity measure is averaged out for 50 comparisons and experiment is repeated for different thresholds of region based matching, varying from 0.3 to 0.8 and the threshold for SIFT matching is kept as constant (0.8). Similarly, the test image is compared with reference images from other classes (dissimilar images) and the similarity measure observed for varied thresholds.



**Figure 7: Averaged region based similarity measure for similar and dissimilar images using varied thresholds for region based matching. The threshold for SIFT matching is set as (a) 0.8 and (b) 0.7.**

Figure 7 (a) shows different plots for three different datasets when test image is compared with similar and dissimilar images. Optimum threshold can be obtained from these plots (empirically) by observing for threshold which gives higher similarity for similar images and lower similarity for dissimilar images. Similar plots (see Figure 7 (b)) are obtained when threshold for SIFT matching is set as 0.7. In both the cases, the optimal threshold is found to be 0.7, which gives maximum separation between similar and dissimilar images. The optimal range of threshold is found to be 0.6 - 0.7.

Experiments were conducted for satellite images exclusively and optimal threshold found out for SIFT as 0.8-0.9 and region based as 0.6 - 0.7. To summarize, the optimal threshold ranges for SIFT matching when using satellite images and real-world optical images are found to be 0.8-0.9 and 0.7-0.8

respectively. The threshold ranges for region based matching for all kind of images is found to be 0.6 to 0.7.

#### **4. CONCLUSIONS**

This paper has proposed a region based SIFT algorithm to provide more meaningful matching between the images. The image is classified into different regions and a fuzzy based method is used to determine the final similarity measure. The proposed method is tested on different optical images, Caltech image datasets and a few satellite images. The method is found to be invariant to rotation, scaling and illumination changes which is inherited from SIFT algorithm and has shown better performance than SIFT. The performance on satellite images can be further improved by adding texture features along with region based approach. Adaptive thresholding is also a scope of future work. The system can be modified and used in variety of applications such as image retrieval, object recognition, recognizing targets in defence applications etc.

#### **5. ACKNOWLEDGMENTS**

The authors acknowledge and thank the POINTS students who have implemented & generated the results.

#### **6. REFERENCES**

- [1] Foschi, P.G., Kolippakkam, D., Liu, H., Mandvikar, A., 2003, Feature extraction for image mining, Joint Conference on Information Sciences (JCIS 2003)
- [2] Addison, J.F.D., Wermter, S., Arevian, G.Z. 2003, A comparison of feature extraction and selection techniques, International Conference on Artificial Neural Networks
- [3] Roth, P.M., Winter, M. 2008, Survey of appearance-based methods for object recognition phd thesis, Inst. for Computer Graphics and Vision, Graz University of Technology, Austria
- [4] Zitova, B., Flusser, J., 2003, Image registration methods: a survey, Image and Vision Computing (2003), 977–1000
- [5] Lowe, D.G. 2004, Distinctive image features from scale-invariant keypoints, Proceedings of the International conference of Computer Vision, Corfu 60 (2004) 91–110
- [6] Park, D.K., Jeon, Y.S., Won, C.S, 2000, Efficient use of local edge histogram descriptor, Proceedings of the 2000 ACM workshops on Multimedia, Los Angeles, California, United States, 51–54
- [7] Squire, D.M., 1998, Learning a similarity-based distance measure for image database organization from human partitionings of an image set, IEEE Workshop on Applications of Computer Vision (WACV98), Princeton, NJ, USA., 88–93
- [8] Caltech, 2012, [http://www.vision.caltech.edu/Image\\_Datasets/Caltech256/](http://www.vision.caltech.edu/Image_Datasets/Caltech256/)
- [9] Wang, 2012, <http://wang.ist.psu.edu/docs/related/>
- [10] NASA, 2012, <http://eol.jsc.nasa.gov/>