J. Basic. Appl. Sci. Res., 2(9)8612-8616, 2012
© 2012, TextRoad Publication

ISSN 2090-4304
Journal of Basic and Applied Scientific Research
www.textroad.com

Objects Recognition Using the Histogram Based on Descriptors of SIFT and SURF

A.Ahmadi[1], Dr. M.R .Daliri[2], Ali Nodehi[3], Amin Qorbani[4]

[1]College of Electricity and Computer, Qazvin Branch, Islamic Azad University, Qazvin, Iran
[2]Department of Electrical Engineering , University of Science and Technology, Tehran, Iran
[3]College of Electricity and Computer, Gorgan Branch, Islamic Azad University, Gorgan, Iran
[4]College of Electricity and Computer, Kordkouy Branch, Islamic Azad University, Kordkouy, Iran

ABSTRACT

This article reviews a method of histograms based on descriptors of SIFT and SURF and their application in Recognition of three-dimensional object recognition from various view. This method will be provided for Recognition of objects in a set of three-dimensional objects and images of ETH80 databases to assess the effectiveness of methods used in determining the three-dimensional objects. Finally, some of these results will be presented.

KEYWORDS : Histogram based on features descriptor, three-dimensional objects recognition , SIFT & SURF, ETH80's images databases.

1. INTRODUCTION

Detection of different objects using image descriptors have grown substantially in recent years and has been too promising in the field of machine vision for having powerful system with extensive and different capabilities so that scientists are more and more hopeful to have systems with similar features of biological vision systems in near future.

According to research by neuroscientists, there is evidence indicating that in the biological visual systems are some neuronal layers which are sensitive to local features and have a major role in process of determination of objects and environmental features. In recent years some local feature extraction algorithms have been proposed which successfully detected the two-dimensional patterns such as face that also have been used in detection and recognition three-dimensional objects, and environment features. SIFT and SURF local descriptors are among most popular descriptors that have been used efficiently in different applications of vision machines. These descriptors in this article are used for detection of three-dimensional objects.

1-1. SIFT descriptors

Lowe(1999) represented Scale Invariant Feature Transform(SIFT) that the objective of the sift algorithm is to extract repeatable feature points from an image. The features are invariant to image scaling, rotation and partially invariant to change in illumination and 3D camera view point. The SIFT feature points are highly distinctive in the sense that a single feature can be correctly matched with high probability against a large database features this provides of basic for object and scene recognition.

Following are the major stage 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.

2. Keypoint localization: At each candidate location, a detailed model is fit to determine location and scale. Keypoints are selected based on measures of their stability.

3. Orientation assignment: One or more orientations are assigned to each keypoint 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.

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.

This approach has been named the Scale Invariant Feature Transform (SIFT), as it transforms image data into scale-invariant coordinates relative to local features.

An important characteristic of this approach is that it generates large numbers of features that densely cover the image over the full range of scales and locations. Usually an image that has 500x500 pixels will have about 2000 features (this features are stable and this number relevant to both image content and choices for various parameters). The quantity of features is particularly important for object recognition, where the ability to detect small objects in cluttered backgrounds requires that at least 3 features be correctly matched from each object for reliable identification.[1]

***Corresponding Author:** A.Ahmadi, College of Electricity and Computer, Qazvin Branch, Islamic Azad University, Qazvin, Iran, Email: aliahmadi@qiau.ac.ir

1-2 **SURF descriptor**

SIFT and SURF algorithms use methods with small differences to detection features. SIFT makes visual pyramids to find candidate points and filters each layer according to the Gauss law with increased Sigma values and finds differences. On the other hand, SURF uses Hessian Matrix to select the candidate points in different sizes and uses, as is used in the Hessian-Laplace method [7].

In the phase of determination the main angle of the feature vector, the angle dominant over vector of feature descriptor is obtained using Haar-Wavelet filters and the integral of the image to speed up the filtering operation, [7].

Finally, using filters Haar-Wavelet the property descriptor is obtained in the related areas to extract property vector in the neighbor of the candidate point like SIFT method but in less dimensions.[7].

1-3 3-D object recognition

Three-dimensional object recognition is one of the charming capabilities of human accompanied him since childhood that often simply and unconsciously is performed by brain. Man in a short glance can recognize a special object that is changed in terms of lighting, angle view, color and different properties. The idealized purpose of machine vision of science is to build a machine vision system that its vision ability is like human or even better. Since the advent of machine vision, methods developed for detection of objects can be classified into three general methods:

1. Methods based on description of the geometric properties
2. Methods based on physical features
3. Methods based on feature descriptors

When objects in the three-dimensional objects are investigated in different angles, then they will have different properties and appearance. Therefore, in order to identify specific object in different angles, a set of structure and appearance properties in different angles are needed. Therefore, three dimensional diagnosis methods are often based on the construction of three-dimensional models mentioned above. However, powerful methods that are used today for various applications of machine vision are mostly used in all three methods to improve the performance and using their benefits to increase power and accuracy of machine vision systems.

This paper presents a method based on feature descriptors that test results will be presented on a database of three-dimensional objects.

2. The suggested method

SIFT and SURF descriptors are used with a direct matching method to track, detect and classify three-dimensional objects as well as two dimensional patterns such as figures. But the method used in this article is similar to the method presented in [8] that will be explained in brief in the next part. Similar objects have the same local structural patterns in their structure. So if you use the Local Identifier like SIFT that describes specific features of an object in a proper way, then you can find two dimensional and three-dimensional models that reflect structural features of related objects that is considered as a criterion to identify and categorize different objects.

The method used in this article first finds SIFT or SURF feature vectors of images related to specific objects and then classifies them using K-Means as a certain number of categories. Then obtained centers of all the objects in question are considered as points between the histogram. Then a descriptor of the histogram between the points will be obtained for any particular object with classification of obtained vectors from feature descriptors of SIFT and SURF. After normalization, the feature vector will be used for classification of images of related objects.

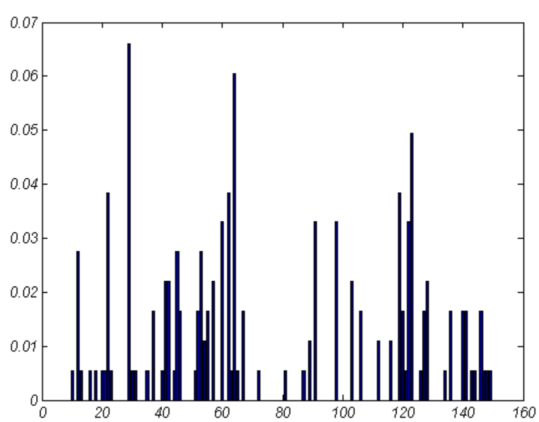

Figure 3: An example of histogram obtained from an image.

Histograms obtained by this method can be classified by one of the different classification methods like neural networks or SVM so that obtained histograms of the different images of each object in different conditions and angles can be put in one group and class. But the suggested method is that histograms or feature vectors corresponding to each object should be considered as points in areas with high-dimensional space that center of these points will be obtained by K - Means method or other methods. Then histogram of the new image related to different objects are obtained and classified. Also Euclidean distance of this point will be obtained via related obtained centers in the previous stage and the closest center is considered as class of the new object. Also other criteria such as distance or Manhattan distance can be used [8] to compare the new histogram with the histogram of the centers and their final categorization.

3. Experiments and results

3-1 Detection of three-dimensional objects from different angles

As previously mentioned, three-dimensional object recognition is one of the most important tasks and capabilities of machine vision science. So this section, as mentioned in [8], provides a method based on databases of images of various objects obtained by the author that is evaluated [9] in different angles. These images consists of two separate images as the training and testing set , each containing 55 images of 11 different objects so that five images of each object is available from a fixed angle, but with different lighting and color conditions. However, test set images with different angles and different lighting conditions of related and machine vision methods should be able to recognize objects from different angles. Due to the above mentioned results obtained by the described, we do not compare this method with other methods of this classification.

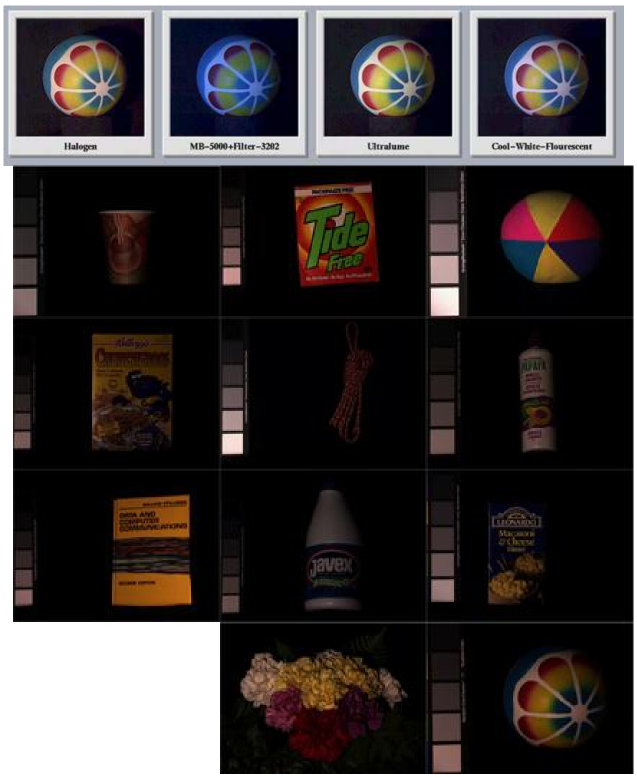

Figure 4: Examples of images of related objects belong to the training set was that are different only in terms of colors and lighting [9].

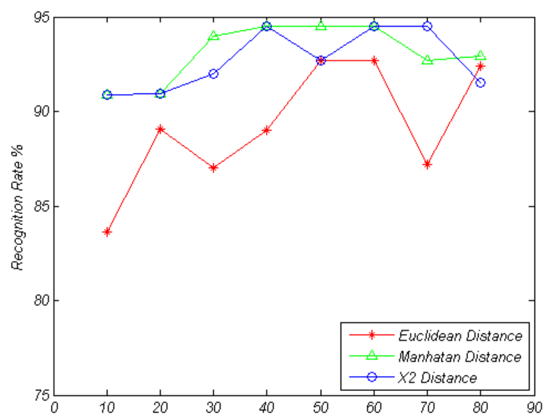

Figure 5: Rate of correct detection of images of objects from the test set depending on the number of classes of histograms obtained by SIFT descriptor of each image.

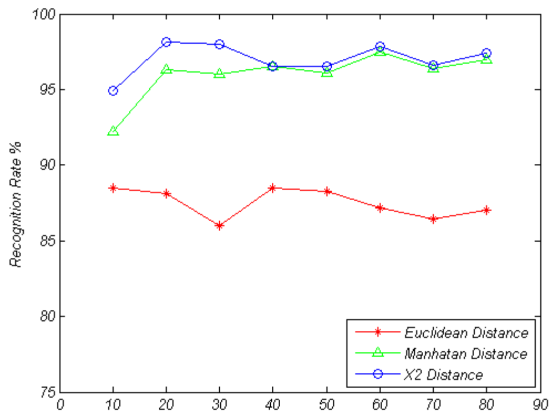

Figure 6: Rate of correct detection of images of objects from the test set depending on the number of classes of histograms obtained by SURF-128 descriptor of each image.

3-2 ***Testing on databases of ETH80***

Next, this method was performed on the standard databases of ETH80 images. Results of the same methods can be observed in the [2]. Some examples of the images can be seen in figure 12.In this collection the database includes 8 topics (classes) of different objects that in each class there are 10 types of that object. There are 41 pictures of each species. In fact, there are 80 objects and 41 images is available from each of object. All objects have been selected so that include different categories of objects in the real world such as fruits ,vegetables ,living creatures ,vehicles and etc.

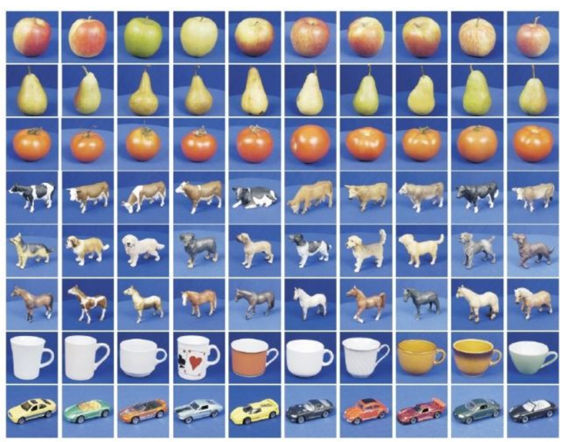

Figure 7 - Sample images of objects from databases of ETH80

The obtained results are given below:

	Leibe Method	SIFT Histogram (X2Dist.)	SIFT Histogram (Manhattan Dist.)
Apple	57.56	69.27	72.68
Pear	66.10	78.54	77.80
Tomato	98.54	72.20	71.46
Cow	86.59	76.09	75.12
Dog	34.63	77.56	78.78
Horse	32.68	70.49	73.90
Cup	79.76	74.39	73.90
Car	62.93	77.56	80.24
Total	**64.85**	**74.51**	**75.49**

Figure 8: Rate of correct detection of the test images depending on various topics in the database of ETH80 images

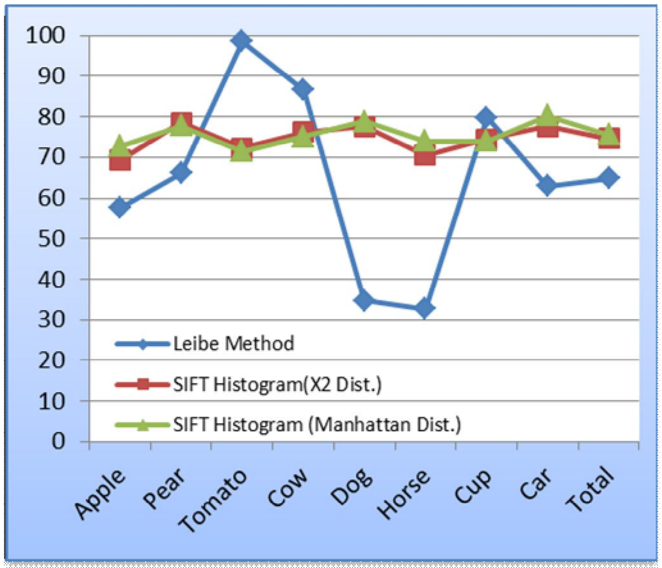

Figure 9: The correct diagnosis rate of 14 images from the test set depending on various topics in the database of ETH80 images

It can be seen that although obtained results in some cases are lower than the results of Leibe & Schiele [2], in the most issues, especially issues with more complex images much better results are obtained. If you use more powerful hardware, the SVM method and genetic algorithm, then it can be expected that better results are obtained.

4. Conclusion

This article investigates a method based on statistical distribution of feature descriptors of SIFT and SURF to detect r two and three dimensional objects from different angles. The obtained the results show the efficiency of this method. But on the other hand with a mix of different descriptors, better performance can be achieved [8]. But for detection of objects that have smooth surfaces, little details or objects without effective erasure, these methods are not very effective. We can combine these methods with methods based on structural and physical properties to give powerful methods for different applications of machine vision.

REFERENCES

[1] D. Lowe. *"Distinctive Image Features from Scale-Invariant Keypoints* ", IJCV, 60 (2) :91-110, 2004.

[2] Leibe.B, Schiele.B, 2003, *"Analyzing Appearance and Contour Based Methods for Object Categorization"*, CVPR'03.

[3] Bay, H,. Tuytelaars, T., & Van Gool, L., 2006, *"SURF: Speeded Up Robust Features"*, 9th European Conference on Computer Vision.

[4] Hetzel.G, 1 Leibe.B, *"3D Object Recognition from Range Images using Local Feature Histograms"*

[5] T. Ojala, M. Pietikäinen, and T. Mäenpää, "*Multiresolution Gray-Scale And Rotation Invariant Texture Classification With Local Binary Patterns*", IEEE Trans. Pattern Analysis and Machine Intelligence, vol.24, no.7, pp.971-987, 2002.

[6] Lowe, D.G. 2001. "*Local feature view clustering for 3D object recognition*". IEEE Conference on Computer Vision and Pattern Recognition, Kauai, Hawaii, pp. 682-688.

[7] Herbert Bay, Andreas Ess, Tinne Tuytelaars, Luc Van,"*Speed Up Robust Features(SURF),*" Elsevier Inc, 2007.

[8] J. Zhang , M. Marszalek, S. Lazebnik, C. Schmid "*Local Features and Kernels for Classification of Texture and Object Categories: A Comprehensive Study,*" International Journal of Computer Vision, Business Media, LLC. Manufactured in the United States, 2006.

[9] Brian Funt , Kobus Barnard , Lindsay Martin, "*Is Machine Colour Constancy Good Enough?,*" In Proceedings the 5th European Conference on Computer Vision, 1998.