Object Recognition using Harris Corner Detection and Neural Network in HSV color Space

Maysam Shahsavari
Department of Electrical Engineering , Qazvin Branch ,Islamic Azad university
Qazvin ,Iran
m.shahsavari@qiau.ac.ir

Mohammad Reza Daliri
Assistant professor of electrical engineering in Science and Technology University,
daliri@iust.ac.ir

Mahyar Haj Abotalebi
Department of Electrical Engineering , Qazvin Branch ,Islamic Azad university
Qazvin ,Iran
mahyar.ab@gmail.com

Abstract

The Harris corner detector is a popular interest point detector due to its strong invariance to: rotation, scale, illumination variation and image noise. The Harris corner detector is based on the local auto-correlation function of a signal; where the local auto-correlation function measures the local changes of the signal with patches shifted by a small amount in different directions. In this paper, with the approach of corner detection system, Harris, we create a specification vector from the entry image that include the color of the pixel from the eight key neighborhoods in the color space, HSV .This specification vector transform to a fifty column histogram and feeds a neural network which will use it for classifying the entry images.

Keywords: HSV, Harris, Neural network, object re Recognition, Corner Detection

1. Introduction

Object recognition is one of the most useful topics in the field of computer vision. The use of local characteristic that is invariant to illumination changes, rotation, and scale can produce good results [1].

In order to do this first, few key points are chosen from the images and then a vector of specifications of those keypoints are produced that can be used to determine similarities between two images.

One of the methods for extracting key points is using sift technique. This technique produces good result. But, for specific object it does not extend well for a group of objects [2,3]. In sift technique an image is converted to set of local specification vectors that are invariant to turning, size change, and limited changes in lighting [4]. Sift technic has a good performance but required high processing power. Harris corner detection method is also one of the method for striking key points is based on determinant and the second degree effect of matrix that can detect corners with high speed accuracy [5]. In this paper harris corner detection method plus nurl network is used to allow system to find images that are learned to the system first. The procedure is that after finding the key points a specification set vector of each image build that is then thought to nural network and eventually this knurl network is used in input image.The advantage of this method is simplicity and high speedup to using to harris method. In this paper 320 images in four categories are use for educating the knurl network. Those images are downloaded from website of " machine vision laboratory " from the department of the department of information technology and electrical engineering of university of ETH[6].In each class they are 80 images.

In this paper the color of points of neighborhood of key points in HSV the color space are used to detect the specification because these colors are less sensitive to changes in light compare to RGB space.

Hsv is defined in a cylindrical space where an the vertical axis defines color, distance from the axis defines saturation and height defines brightness.

2. Harris

The formula for discovering the key point is as follows:

$$R = Det(M) - k \times Tr(M)^2 \quad (1)$$

Where K is a predefined parameter and M is the second degree effect of this matrix:

$$(2)\quad M = \begin{bmatrix} \left(\frac{\partial I}{\partial x}\right)^2 \otimes w & \left(\frac{\partial I}{\partial x}\right)\left(\frac{\partial I}{\partial y}\right) \otimes w \\ \left(\frac{\partial I}{\partial x}\right)\left(\frac{\partial I}{\partial y}\right) \otimes w & \left(\frac{\partial I}{\partial y}\right)^2 \otimes w \end{bmatrix}$$

W is the weight of the gaussian filter and I is a function of the brightness of the photo. ⊗ denotes convolution. This operator is obtained from changing the average of grey surface is a square window that slides between two vertical directions. Corners are optimums in Harris messier with a threshold more than a certain value.

$$\{x_c\} \quad \{x_c \mid R(x_c) > R(x_i), \forall x_i \subset w(x_c)\} \qquad (3)$$
$$R(x_c) > t_{threshold}$$

{xc} is a set of all corners R{x} is the Harris messier at x, W(xc) is the 8 neighborhoods around Xc ,and threshld is the given threshold.

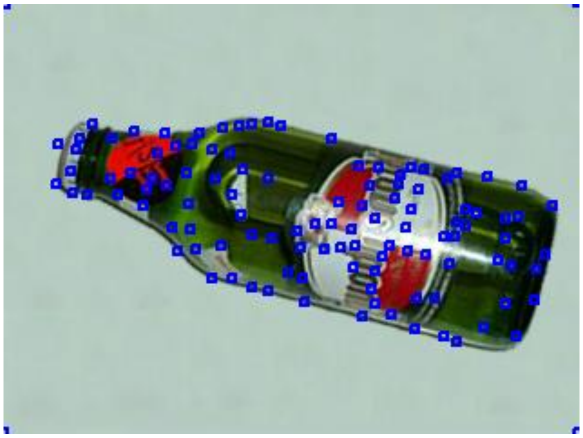

Figure 1: One of the photos for which crier detection is done using Harris method. Blue dots are the extracted key points

3. HSV Color Space

HSV color space is defined in a cylinder space where color changes with respect to angle, such that red is angle 0, green is angle 120, and blue is angle 240. Vertical axis represents shades of gray. Bottom of the axis is black and top of the axis is white. Distance from center defines saturation.

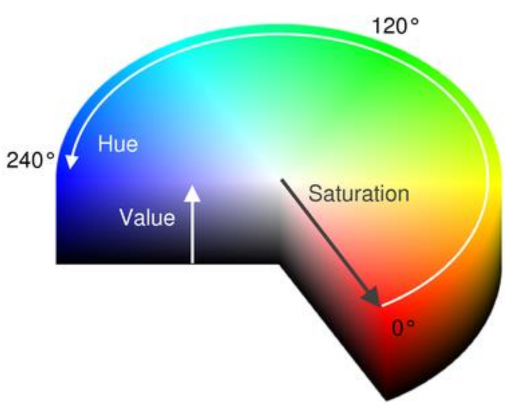

Figure 2: HSV cylinder

4. Specification Vector

As mentioned in the introduction section, when key points are used, photo is converted to a set of specification vectors. in this article we form a matrix of HUe characteristic using 8 neighborhoods of the key points extracted using Harris method. Since our goal was to train specific number of inputs, our matrix did not have a constant number of enteries because each image has a deferent number of key points. To solve this with this we created a 50 column histogram the matrix that devised the color range into 50 segments and had a similar format for every class of images.

The necessarily time for extracting the histogram from beginning to the end of operation was between 300 to 350 mile second. Using an intel processor model core is 720 QM. Using implemented sift method [8] the key point for two of the photos used in this paper were generated. This took about 1.4 second with the same computer.

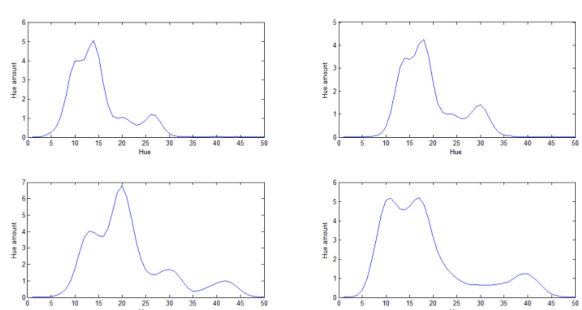

Figure3:four examples of class A specification vectors

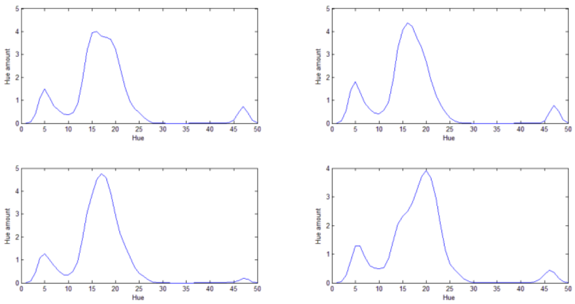

Figure 4 : four example of class B specification vectors

As pointed out before, the specification vector used in this article includes 50 columns of colors of the points of 8 neighborhoods of the key points, which are extracted using Harris method. Based on observations of each photo, 2200 to 2400 points where extracted and for sampling purposes we need square root number of this points or about 50.

Photos that we used in our study all have simple backgrounds and the target object is the only one in the image. Images are from different angles, rotated, slightly magnified, and in different light conditions.

Figure 5: a sample of used class

5. Neural Networks

After obtaining the specification vector we need to train a neural network so that it can recognize an input image. Nural network is a large parallel processor that intends to store experimental information and use them[9]. One of the algorithms employed for multi-layer N.Ns is the post-error propagation algorithm that runs in two stages. In forward stage the free parameters of the network are initialed. Then the error values are calculated from equation 4:

$$e_i = d_i - y_i \qquad (4)$$

Where di is expected output and yi is the actual output.

In the second stage, the error value ei is propagated back in the network to turn the error values [10,11].

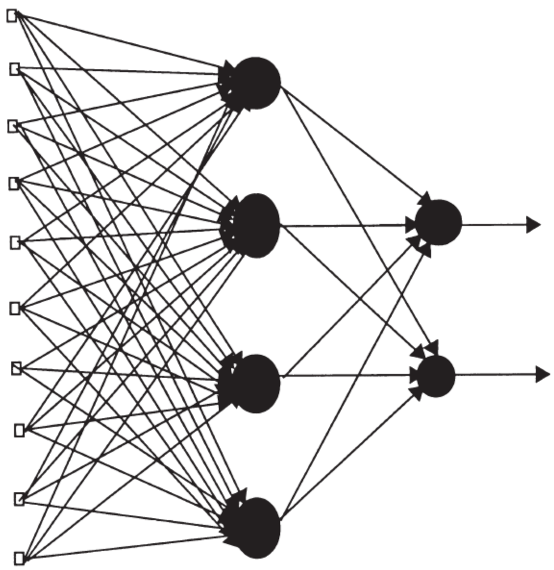

Figure 6: The structure of a multilayer network.

The implementation in this paper uses a neural network with a hidden layer. The input matrix has 320 columns and 50 rows. 70% of this values are used for training, 15% for testing, and 15% for validation. The number of neurons in the neural network are determined by trial and error[12]. The table below shows the average detection by the neural network for 4 different values.

Number of neurons	80 *neurons*	70 *neurons*	65 *neurons*	60 *neurons*
Error in network	25.7%	13.5%	6.42%	19.5%
deviation	4.58	1.67	2.1	17.

Table 1: based on the table, 65 was chosen for the number of neurons in the hidden layer to minimize error.

6. Conclusion

In this paper we presented a fast method for classification of images using neural networks and Harris corner detection. this method is useful specially in applications with limited number of input images and classes do not have a lot of color similarities.

Using HSV color spaces reduces sensitivity to lighting variations and Harris corner detection increases the speed of producing specification vectors. To improve this solution this solutions one can use object recognition algorithms so that the input images are not restricted to those with simple background.

References

[1] H.Bay, A. Ess, T. Tuytelaar, and L. Van Gool: "Speeded-Up Robust Features (SURF)", *Computer Vision and Image Understanding 110(3):* 346-359, 2008.

[2] S. Batra, 'Multi-Class Object Recognition Using Shared SIFT Features", Learning Report, Machine Learning Cs229, 2007.

[3] David G. Lowe, "Object recognition from local scale-invariant features" International Conference on Computer Vision, Corfu, Greece (September 1999).

[4] X. Wu, Z. Shi, Y. Zhong, "Detailed Analysis and Evaluation of Keypoint Extraction Methods" International Conference on Computer Application and System Modeling, 2010.

[5] Harris, C., Stephens, M. "A Combined Corner and Edge Detector" Proceedings

of 4th Alvey Vision Conference, Manchester, pp. 147-151, 1998.

[6] http://www.vision.ee.ethz.ch/datasets/index.en.html

[7] Alvy Ray Smith “Color gamut transform pairs,” Computer Graphics,12(3):12-25, 1978.

[8] Demo code for detecting and matching SIFT features, http://www.cs.ubc.ca/~lowe/keypoints/

[9] Haykin S.,” Neural networks: A comprehensive foundation, 2nd ed. “, Prentice-Hall.,1998.

[10] Werbos, P. J “Beyond Regression: New Tools for Prediction and Analysis in the Behavioral Sciences” PhD thesis, Harvard University, 1974.

[11] Rumelhart, D. E., Hinton, G. E., and Williams, R. J. “Learning internal representations by error propagation. In Parallel Distributed Processing”, volume 1, pages 318-362. MIT Press, 1986.

[12] Shin-ike, K. “A two phase method for determining the number of neurons in the hidden layer of a 3-layer neural network” SICE Annual Conference, 2010.