The Comparison of Object Recognition and Identification by Using Image Processing Base on the Neural Network, the Hough transform and the Harris Corner Detection

Jaruwan Toontham, Chaiyapon Thongchaisuratkul

Abstract—This paper presents an object recognition and identification system using the back propagation neural networks. The performance of Hough Transform and the Harris Corner detection are compared with the following procedures and methods; the webcam is used to capture the object and create an input image, change the color image from RGB to gray scale, resize, learn and recognize the objects by neural network, and separate the objects by the robot arm. Three different types of objects in this study are triangle, rectangle and rigid circle. The object recognition and identification from the neural network, the Hough transform, and the Harris corner detection are compared. The results showed that the neural network gives more accuracy than the Hough transform and the Harris corner detection.

Index Terms—Hough Transform, Harris Corner Detection, Back Propagation Neural Network

I. INTRODUCTION

The Object recognizing is one of the most challenging problems in computer vision [1]. The combination of computer vision and machine learning techniques has recently been considerable progress in developing real world object recognition system base on the use of invariant local feature [2]—[4]. In this research, the neural network is applied for an object recognition and identification. The Hough transform and the Harris corner detection are studied for comparing with neural network. There are many research studies employed the neural network for different reasons. Back propagation algorithm was used for recognize and identify the construction of Batu Aceh [5]. Fourier Descriptor Pattern was used to classify the objects resulting as the vectors. The back propagation algorithm used these vectors to recognize and identify the objects was introduced by Qureshi and Jalil [6]. Multi-layer neural networks including forward and back propagation for real time object recognition is illustrated for teaching [8]. The Hopfield Neural Network is used to recognize the parts of the computer motherboard. The result of this technique was simulated in Matlab and discussed in [9].

In our work, a robot arm was designed to use a neural network method to recognize, identify, and separate the objects. The triangle, rectangular, and rigid circle objects are used to verify the system.

II. IMAGE PREPROCESSING AND ENHANCE

The images of the object from webcam that used in our experiment have different quality such as brightness, contrast, or noise. The fuzzy image processing (FIP) is used to adjust brightness of the image. The process includes image fuzzification, membership modification, and image defuzzification which can be showed in Figure 1 [9].

![Fig. 1 Fuzzy Image Processing](image)

To adjust the brightness of the image, the INT operator with the following methods is used. The original image matrix \( X \) of \( N \times M \) is the array of fuzzy singleton of each member represented by the degree of brightness levels \( p, p = 0,1,2,3, \ldots, P - 1 \) with \( p \) values between 0 and 255 can be written as follows,

\[
X = \begin{bmatrix}
\mu_{11}/x_{11} & \mu_{12}/x_{12} & \ldots & \mu_{1M}/x_{1M} \\
\mu_{21}/x_{21} & \mu_{22}/x_{22} & \ldots & \mu_{2M}/x_{2M} \\
\vdots & \vdots & \ddots & \vdots \\
\mu_{N1}/x_{N1} & \mu_{N2}/x_{N2} & \ldots & \mu_{NM}/x_{NM}
\end{bmatrix}
\]

where \( 0 \leq \mu_{nm} \leq 1, m = 1,2, \ldots, M, n = 1,2, \ldots, N \). The equation (1) is used to calculate.

\[
T_i(\mu_{\text{new}}) = T_i(\mu_{\text{old}}) = 2\mu_{\text{old}}^2, \quad 0 \leq \mu_{\text{old}} < 0.5
\]

\[
= T_i(\mu_{\text{old}}) = 1 - 2(1 - \mu_{\text{old}})^2, \quad 0.5 \leq \mu_{\text{old}} \leq 1
\]

Each pixel of image brightness is adjusted as follows, if the values are greater than 0.5 increases the brightness of the pixels, if the values are less than 0.5 decreases the brightness.
of the pixels. The procedure of the fuzzy image enhancement is showed in Figure 2.

![Fuzzy Image Enhancements](image)

**Fig. 2 Fuzzy Image Enhancements**

III. EASE OF USE PIXEL AVERAGE MEAN

The pixel average mean is used for resize before the learning process.

A. **Change the image color**

The image colors are changed to gray scale.

B. **Resize the image**

The image are resize to $14 \times 32$ pixel.

C. **Calculate pixel average mean**

The $y$ axis of the image is calculated for pixel average mean. The result has 32 columns per image. The example pixel of the gray scale showed in Figure 3.

```plaintext
<table>
<thead>
<tr>
<th>120 35 80 120 62 120</th>
</tr>
</thead>
<tbody>
<tr>
<td>65 35 80 120 62 20</td>
</tr>
<tr>
<td>120 35 80 120 62 120</td>
</tr>
<tr>
<td>120 35 80 120 62 120</td>
</tr>
<tr>
<td>203 35 80 120 62 255</td>
</tr>
<tr>
<td>120 35 80 120 62 120</td>
</tr>
<tr>
<td>100 35 80 69 62 20</td>
</tr>
<tr>
<td>255 35 80 120 62 200</td>
</tr>
<tr>
<td>120 35 80 14 62 2</td>
</tr>
<tr>
<td>255 255 255 120 62 120</td>
</tr>
<tr>
<td>120 35 80 120 62 120</td>
</tr>
<tr>
<td>120 35 80 120 62 120</td>
</tr>
<tr>
<td>120 35 80 120 62 120</td>
</tr>
</tbody>
</table>
```

**Fig. 3** The example pixel of gray scale.

IV. THE NEURAL NETWORK

Artificial Neural Networks are created in the computer as the simulation. The results are illustrated that how the computer works as the human brain for pattern recognition, listening, or reading the human language. This can be implemented as a "robot" [10], [11].

The structure of neural network consists of input units, hidden units, and output units. The paths are connected between each unit have weights which can be changed during learning phase. Neural network learning is based on back propagation algorithm. The neural network learning techniques are considered as a human learning which can be identified as supervised and unsupervised learning. Supervised learning means learned by taught [12] such as the students are learning by the teachers. Unsupervised learning means learned without taught such as the student can be identified the plant and animal species by shape or by behavior without teaching [13].

In this research, two neural network methods—Learning Vector Quantization Neural Network (LVQ)[14] and Back Propagation Neural Network (BPN) are implemented. The method that gives more accuracy is selected.

A. **Learning Vector Quantization Neural Network**

The structure of Learning Vector Quantization neural network shown in Figure 4 consists of 32 inputs, 10 competitive layers and 3 outputs.

![LVQ Network](image)

**Fig. 4** The structure of Learning Vector Quantization Neural Network

LVQ has 3 layers including input layer, Kohonen classification (competitive) layer, and output layer.

![LVQ Performance](image)

**Fig. 5** The mean square error and epoch of LVQ

B. **Back Propagation Neural Network**

The structure of back propagation neural network showed in Figure 6. Experimentations for finding the best method of performance and gradient showed in Table 1.
TABLE 1 PERFORMANCE AND GRADIENT IN VARIOUS FORMATS.

<table>
<thead>
<tr>
<th>Format</th>
<th>Performance</th>
<th>Gradient</th>
<th>Train</th>
</tr>
</thead>
<tbody>
<tr>
<td>32-10-10-1</td>
<td>4.58e-5</td>
<td>0.02</td>
<td>Trainlm</td>
</tr>
<tr>
<td>32-10-20-1</td>
<td>0.000785</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>32-30-30-1</td>
<td>1.77e-5</td>
<td>0.0186</td>
<td></td>
</tr>
<tr>
<td>32-10-10-1</td>
<td>0.001</td>
<td>0.00523</td>
<td>traingd</td>
</tr>
<tr>
<td>32-20-20-1</td>
<td>0.001</td>
<td>0.00583</td>
<td></td>
</tr>
<tr>
<td>32-30-30-1</td>
<td>0.001</td>
<td>0.00655</td>
<td></td>
</tr>
<tr>
<td>32-10-10-1</td>
<td>0.001</td>
<td>0.00125</td>
<td>traingdm</td>
</tr>
<tr>
<td>32-20-20-1</td>
<td>0.001</td>
<td>0.00736</td>
<td></td>
</tr>
<tr>
<td>32-30-30-1</td>
<td>0.516</td>
<td>1.33</td>
<td></td>
</tr>
</tbody>
</table>

The best format is selected by gradient 0.00125, because the value was close to setting target. The threshold value is 0.0001 for 32 input layers, 10 hidden layers and 1 output.

The accurate percentage comparison between LVQ and BPN found that the BPN gets accurate percentage 97.54 more than the LVQ which gets accurate percentage 74.5455. Therefore, in this study the supervisor BPN is selected to learn and recognize objects.

V. CONVENTIONAL HOUGH TRANSFORM (CHT)

The conventional Hough transform is used to recognize the shape by detecting linear curve and circle in the picture [15,16,17,18]. For detection the linear line, the equation (2) is given.

$$\rho = x \cos \theta + y \sin \theta$$  \hspace{1cm} (2)

If the $x, y$ coordinate of the pixel on the image plane $x, y$, $\rho$ is perpendicular, the distance is the origin to straight lines in plane image $x, y$, the $\theta$ is the angle from $x$-axis perpendicular to $y$-axis.

Two-dimensional accumulator cell $A(i, j)$ was created which the components of $\rho, \theta$, for every pixel with coordinates $x, y$ images of any pixels on the same line in the same accumulator cell. When the $\rho$ and the $\theta$ have the same amount, and the accumulation in the accumulator cell is greater than the specified threshold, the values are converted back to the coordinates $x, y$. This recognizes that the line is rated at point on the plane $x, y$. In the same way, the CHT can find a curve and a circle as given in equation (3).

$$r^2 = (x - C_x)^2 + (y - C_y)^2$$  \hspace{1cm} (3)

When the point $(C_x, C_y)$ is the center of the circle along the $x$ and $y$ axis, $r$ is the radius of the circle or arc. In the case of a circle, accumulator cell is $A(i, j)$ 3-dimensional, the parameters contain $C_x$, $C_y$ and $r$. These parameters are used
VI. HARRIS CORNER DETECTION

The Harris Method is used for object corner detection. This method does not change the value of scale, noise, rotation, variance of light by calculating the local auto correlation function [2], [19]—[21]. This function considers a signal measures from the local changes of the signal with patches shifted by a small amount in different directions as given in equation (4).

\[ c(x, y) = \sum [I(x_i + \Delta x, y_j + \Delta y) - I(x_i, y_j)]^2 \]  

(4)

When \( w(x, y) \) is a window function, \( I(x + u, y + v) \) is a shifted intensity, and \( I(x + y) \) is an intensity. The small shifted windows are approximated by equation (5).

\[ E(u, v) \approx M \begin{bmatrix} u \\ v \end{bmatrix} \]  

(5)

\( M \) is a 2x2 matrix which can be calculated by the image derivatives as shown in equation (6).

\[ M = \sum w(x,y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \]  

(6)

If window was shifted, the intensity value and the eigenvalue of \( M (\lambda_1, \lambda_2) \) are changed. There are three cases to be considered:

1) If both \( \lambda_1, \lambda_2 \) are small, so that the local auto-correlation function is flat, the windowed image region is an approximately constant intensity.

2) If one eigenvalue is high and the other low, so the local auto-correlation function is ridge shaped, then only local shifts in one direction (along the ridge) cause little change in \( C(x, y) \) and significant change in the orthogonal direction, this indicates an edge.

3) If both eigenvalues are high, so the local auto-correlation function is sharply peaked, then shifts in any direction will result in a significant increase, this indicates a corner.

VII. CONVEYER SYSTEM AND ARM

The conveyor system simulation was designed using a servo motor driving the conveyor belt 10 cm width, 62 cm long, as shown in Figure 8. The robot arm is a revolute robot type (RRR) and controlled by the microcontroller as shown in Figure 9.

VIII. EXPERIMENTS AND RESULT

The aim of this study was to compare the efficiency of inspection triangle, squares, and circle objects by using neural network, the Hough transform, and the Harris Corner Detection. The software was developed by using the Opencv library and Microsoft Visual C++. The results are collected by testing three different objects. The revolute robot arm was used to catch the object for separating. There are two experiments were implemented.

A. Each object type is implemented for robot to detect and separate 100 times. The result is showed in Table 2.

B. Three object types are mixed for robot to detect and separate. This experiment consists of 100 pieces of each object type. The result is shown in Table 3.

<table>
<thead>
<tr>
<th>Object</th>
<th>Hough Transform</th>
<th>Corner Detection</th>
<th>BPN</th>
</tr>
</thead>
<tbody>
<tr>
<td>rigid circle</td>
<td>98</td>
<td>99</td>
<td>99</td>
</tr>
<tr>
<td>Rectangular</td>
<td>97</td>
<td>95</td>
<td>99</td>
</tr>
<tr>
<td>Triangular</td>
<td>96</td>
<td>95</td>
<td>99</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Object</th>
<th>Hough Transform</th>
<th>Corner Detection</th>
<th>BPN</th>
</tr>
</thead>
<tbody>
<tr>
<td>rigid circle</td>
<td>98</td>
<td>99</td>
<td>99</td>
</tr>
<tr>
<td>Rectangular</td>
<td>96</td>
<td>95</td>
<td>99</td>
</tr>
<tr>
<td>Triangular</td>
<td>96</td>
<td>94</td>
<td>99</td>
</tr>
</tbody>
</table>

The performance of both Harris corner detectors and Hough transform were reduced while the objects have more corner points. The percentages of detect and separate the same objects and the three objects are similar results. The
experimental results of black propagation neural network get more accurate than the Hough transform and the Harris corner detection.

IX. CONCLUSIONS AND DISCUSSION

The black propagation neural network can be used for the object identification and shape recognition. The algorithm can be implemented with the robot arm to separate the objects. The experimental results get more accurate than the Hough transform and the Harris corner detection. The error is occurred form the light affect and the angle of the camera. However, the error for black propagation neural network can be occurred since the training data was difference. The maximum error was 1% for the black propagation neural network.

REFERENCES


Jarawan Toontham received her B.S.Ed. from Rajamangala University of Technology, Thailand, in 1994. She received her MS.Ed from the King Mongkut's University of Technology North Bangkok (KMUTNB), Thailand, in 1999. Now, she is a Ph.D student in King Mongkut's University of Technology North Bangkok (KMUTNB), Thailand.

Chaiyapon Thongchaissuratkrul, Ph.D. He received his B.S.Ed. and M.Eng from the King Mongkut's University of Technology North Bangkok (KMUTNB), Thailand, in 1989 and 1992 respectively. He received his Ph.D. from the Vanderbilt University, USA in 2001. He joined KMUTNB as a lecture of Teacher Training in Electrical Engineering Department in 1989. Now he is an assistant professor in the department. He served as the vice president for network and conference of the Thai Robotics Society in 2008. His research interests include intelligent control systems, robotics, energy conservation, and energy resources.