Image Recognition of Household Objects Based on Shape and Haar-like Features

Jong-Hann Jean[#1], Ming-Jin Hsieh[#2], Kuo-Tsung Tseng[*3], heng-Wei Lin[#4]

[#]*Department of Electrical Engineering, St. John's University*
499, Sec. 4, Tam King Road Tamsui, Taipei, Taiwan R.O.C
[1,2,4]`jhjean@mail.sju.edu.tw`

[*] *Mechanical and Systems Research Laboratories , Industrial Technology Research Institute, Taiwan*
195, Sec. 4, Chung Hsing Rd., Chutung, Hsinchu, Taiwan 31040, R.O.C.
[3] `kttseng@itri.org.tw`

***Abstract*— In this paper we present some preliminary results of our project concerning about developing an image recognition technique for detection of certain household objects, mainly based on the shape and the Haar-like features. Shape feature is used to find the position of certain shape household objects preliminarily. Then, Haar-like feature-based recognition may be only performed for small regions around the positions of detected shape features. Thus, the proposed two-stage recognition can reduce the computation. In addition, experimental results are also presented to prove that the fusion of the two feature detection strategy will complement each other and improve the detection rate and the robustness.**

***Keywords*— image recognition, shape detection, Adaboost, random hough transform**

1. INTRODUCTION

Recently, development of intelligent robot systems to provide housekeeping and home care services has attracted the attention of more and more researchers. For example, Mitsubishi Heavy Industries, Ltd. has developed the Wakamaru robot [1] as a domestic robot that can recognize faces and make autonomous movements without human assistance. Toshiba Corporation has created an updated version of the prototype housekeeping robot, ApriAttenda [2], which can open the door of a refrigerator and picks up a bottle. In addition, University of Tokyo researchers have designed a series of robots [3] capable of performing a variety of housekeeping jobs, like dish-washing and floor- sweeping.

Besides, the intelligent robot systems have been further employed in the field of smart homes for their capability to equip with a variety of sensors and actuators to percept, monitor and control the home environment. For example, the Farglory Realty in Taiwan has brought up a perspective to incorporate housekeeping robot services in smart homes.

The housekeeping or care robot may partially have to operate in autonomous mode in order to accomplish the housekeeping duties. Accordingly, the robot may need to build up an understanding of household objects in home environment through a variety of sensors, among which vision sensor is the most powerful one to distinguish household objects.

In this paper we present some preliminary results of our project to design image recognition techniques for detection of certain household objects in home environments, which can provide housekeeping robots the capability of target identification and self-localization. We first investigate an object detection scheme using the object shapes, which provides first-stage classifications of household objects. Then, we apply a fast object recognition algorithm based on Haar-like features for exact identification of household objects. We integrate the two completely different approaches to improve the computational efficiency of the object recognition process and increase the detection rate or reduce the miss rate.

The remainder of the paper is organized as follows. Section II describes the experiment setup and the proposed image recognition. Section III introduces the method using RHT to detect ellipse and Section IV introduces the detection of another shape, rectangle, based on contour information. Classification based on haar-like features is described in Section V. In Section VI, we present some experimental results and analysis to prove our project. Final, conclusions are given in Section VII.

2. EXPERIMENT SETUP

The scenario of our project is considered to utilize a PTZ camera mounted on a mobile robot to detect target objects in 3-D home environment. For this purpose, we first implemented a shape detection algorithm which is efficient to support real-time image processing and less affected by the noise, background, and occlusion. The idea is to apply the shape detection algorithm to detect and classify household objects preliminarily. Then, utilize the recognition algorithm based on Haar-like features, which substitute the complex process of target pattern matching with inequality tests on simple weighted sum of rectangle regions, to identify the household objects.

The experiment setup is shown in Figure 1, where a PTZ camera is built over a robot platform (1.1m in height) to capture images and the target object is put on the table (0.73m in height) about 0.5m~1.5m away from the robot.

In training phase, we collect training samples of household objects by manually choosing samples of size more than 20x20 pixels but less than 100x100 pixels from captured images, then applying geometric transformation to the captured images for simulation of the robot's views in different angles and different distances. Thus, the training samples are used to define dynamic range of shape parameters for use in shape detection and train the Haar-like features, which form the classifier in the second stage for recognizing the household objects.

In recognition phase, the flowchart of the overall image recognition process is shown in Figure 2 and illustrated as follows. First, the robot performs the shape detection algorithm to detect whether there is an ellipse or rectangle shape in the captured image and then check whether the shape parameters, e.g. the ratio of the major to the minor axis for detected ellipses or the ratio of the width to the length for detected rectangles, are within the assumed dynamic range of household objects. If yes, it further performs the recognition process based on Haar-like features locally around the detected shapes and determine what the household object is (clock or computer monitor in the paper). Otherwise, it will apply the Haar-like feature-based recognition to the whole image to directly find out household objects. When detected, it further performs the shape detection algorithm locally to verify the result.

The proposed two-stage recognition can reduce the computation cost because the Haar-like feature-based recognition may be only performed for small regions around the positions of detected shape features. Only if there is no ellipse or rectangle shape detected, the recognition will be performed for the whole image. In addition, fusion of the two feature detection strategy will complement each other and improve the detection rate and the robustness. The detail of the two feature detection methods are discussed in the following sections.

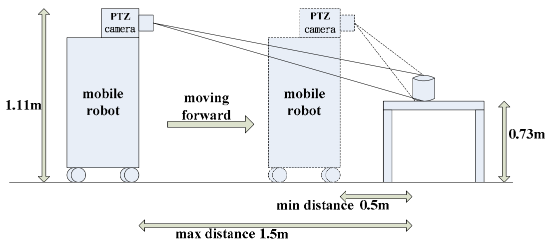

Fig. 1 Experimental environment

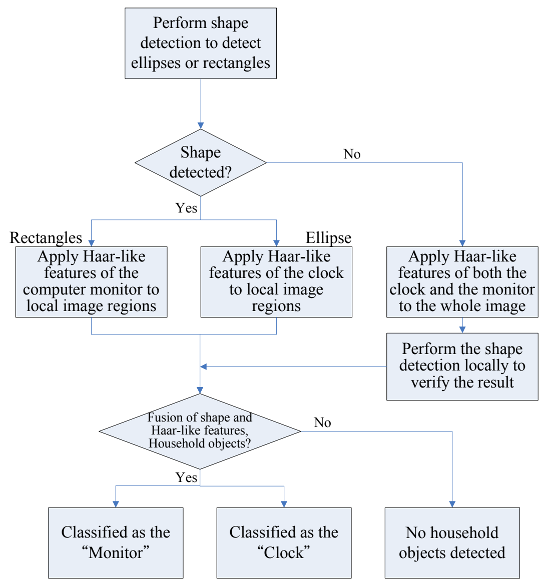

Fig. 2 The flowchart of image recognition process of household objects

3. ELLIPSE DETECTION USING RHT METHOD

The Hough transform (HT) is a standard technique for detecting parametric shapes. But, the basic Hough transform method suffers the problem in performance [4]. The higher dimension of the parameter space, the more amount of memory and computation time is required. To overcome the performance issue the RHT method was first introduced in 1990 by Xu et al. [4] and was successfully extended by

McLaughlin [5] to detect shapes with nonlinear parametric equations.

In comparison to the HT method, the RHT method maps a randomly sampled n-tuple of image pixels to a candidate parameter and records the parameter together with a count into the accumulator array. Candidate parameters with the count over a predefined level indicate the curves present in the image. Because only a small random subset of image pixels is selected, the RHT method reduces the memory and computation time requirements needed to detect curves in an image.

Based on different parametric equations, the RHT method can be used to detect various analytic shapes such as circles, ellipses, triangles, and polygons. In the following we demonstrate the shape detection algorithms for the representative geometric primitives, ellipse. It is easy to extend the strategy to track circles, polygons and other parameterized shapes.

In Algorithm 1, we describe the proposed ellipse detection method. We implement the RHT scheme with reference to McLaughlin's work [5] and Schuler work [6] and use an efficient least-squared fitting method [7] to compute the ellipse parameters.

Algorithm 1: Ellipse Detection Based on RHT

Step 1: Select five pixels randomly from the edge pixels of the image supposing these pixels are located on an ellipse.

Step 2: The classic least-squares fitting technique is used to determine the coefficients *a* to *f* of the ellipse equation as shown below.

$$\begin{aligned} &ax^2 + bxy + cy^2 + dx + ey + f = 0, \\ &4ac - b^2 > 0. \end{aligned} \tag{1}$$

Step 3: Check the inequality $4ac-b^2>0$, which has to be true for a valid ellipse. If false, go to Step 1.

Step 4: Compute the 5-tuple parameter (x_e, y_e, r_a, r_b, θ) which consists of the ellipse centre, the radii of the major and minor axes, and the angle of the major axis, defined by the following equation.

$$\begin{aligned} &\frac{((x-x_e)\cos\theta + (y-y_e)\sin\theta)^2}{r_a{}^2} \\ &+ \frac{((y-y_e)\cos\theta - (x-x_e)\sin\theta)^2}{r_b{}^2} = 1 \end{aligned} \tag{2}$$

Step 5: Search already found parameters in the accumulator array and compare them with the newly found parameter. If there is no equal one, insert the new parameter together with a count of value 1 (totally 6-tuple) into the accumulator array, else simply increase the count of the existent parameter by one.

Step 6: After a specified amount of parameters is found, select parameters with the count value over a predefined level for further check.

Step 7: Perform the shape matching as follows to further check if the found parameters are valid. For each found parameter, determine an ellipse. Grow the ellipse to a constant width and form an elliptical ring. Find and count the edge pixels which lie within the elliptical ring. If the counting result is greater than a predefined threshold, indicate that the ellipse is actually existent.

4. Rectangle Detection Based on Contour Information

When recognizing household objects with rectangle shape, one may keep in mind with the features that the rectangle-shaped object has four certain vertexes, four right angles, the ratio of width to length, and the area of the target object. By estimating these features, the object with rectangle-like shape can be recognized successfully in the image.

Algorithm 2: Rectangle Detection

Step 1: Find the contours having four vertexes from the edge points of the captured image.

Step 2: Check the four angles if they all approximate 90 degrees. When all above steps are satisfied, the shape of the object is rectangle certainly.

Step 3: Estimate the ratio of width to length from the contour having four vertexes, which are used to classify if the object shape is similar to a household object (e.g. computer monitor).

5. Classification Based on Haar-Like Features

The research applies the Haar-like feature-based recognition technique to classification of household objects, which uses a cascade structure of simple classifiers based on simple features, learned through the AdaBoost algorithm. The technique is first proposed by Viola [8] and can achieve fast object detection by utilizing simple Haar-like features and an image representation efficient for computation. In addition, by cascading a series of Haar-like feature-based classifiers, it dramatically increases the detection rate. Its working principle is illustrated as follows.

5.1. Haar-like Features

Using simple features to recognize objects simplifies the computation complexity of image processing so that it speeds up the recognition process. The simple features adopted in the research are associated with the Haar-basis function which has been proposed by Papageorgiou [9]. A typical rectangle Haar-like feature, as shown in Figure 3, is composed of a set of parameters, $\gamma = (x, y, w, h, \alpha)$, which must satisfy the following inequalities: $0 \le x, x+w \le W$, $0 \le y, y+h \le H$, $x, y \ge 0$, $w, h > 0$, $\alpha \in \{0°, 45°\}$, where W and H are width and height of the test window.

Figure 4 shows various types of Haar-like features, each of which is composed of two rectangle regions at most. For such feature type, Viola [8] proposes a new image representation, namely the integral image, and an efficient algorithm to compute it rapidly. Thus, the computation can be quickly completed, even there are mass of candidate features.

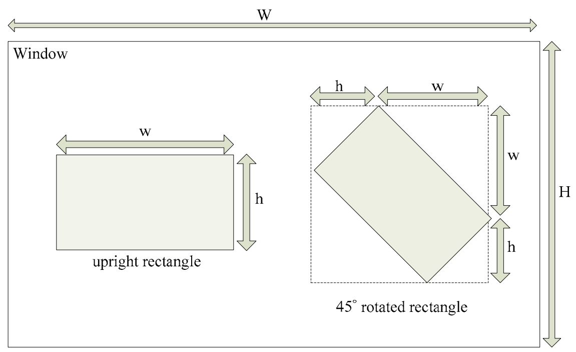

Fig. 3 Examples of upright and 45° rotated rectangle Haar-like features

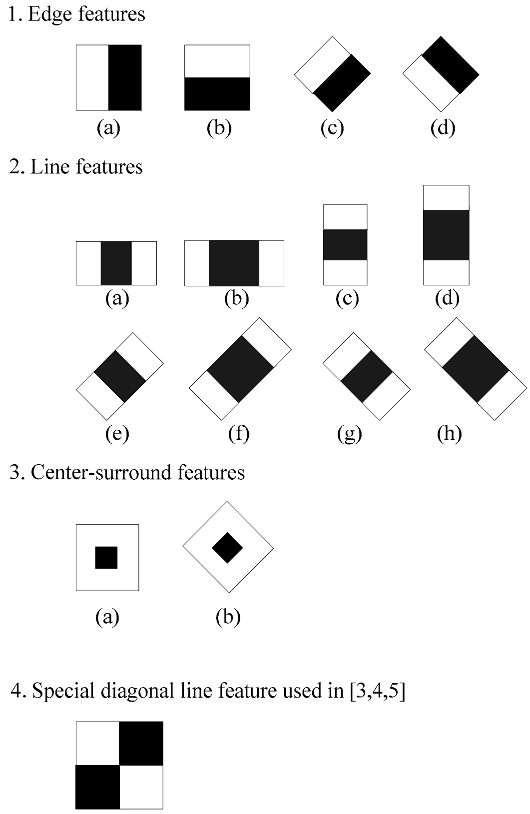

Fig. 4 Types of Haar-like rectangle features

5.2. AdaBoost Classifier

Boosting is a learning algorithm to combine several weak classifiers into a strong classifier. Because the performance of each weak classifier is only required a little better than the expected value, it can be designed simply and computed efficiently. Among various boosting algorithms, Viola [8] adopted AdaBoost algorithm which iteratively re-weights the weights of the weak classifiers based on their classification errors evaluated on the probability distribution of samples in the training set. The probability distribution of samples in the training set is also iteratively updated based on whether the classification on each sample is correct and the classification error rate on the whole training set. One version of the AdaBoost algorithm is described as follows.

Algorithm 3: Discrete Adaboost [8]

1. Given sample images $(x_1, y_1),\ldots,(x_n, y_n)$ where $y_i = \{-1, 1\}$ for negative and positive samples respectively, initialize the weights for samples as

$$w_{1,i} = \{\frac{1}{2m}, \frac{1}{2l}\}, \tag{3}$$

where m and l are the number of positive samples and the number of negative samples respectively.

2. For $t=1,\ldots,T$ (maximum number of weak classifiers):
   a. Normalize the weights,

$$w_{t,i} = \frac{w_{t,i}}{\sum_{j=1}^{n} w_{t,j}}, \tag{4}$$

   where w_t can be viewed as the probability distribution of samples in training set.

   b. For each feature, j, train a classifier h_j which is restricted to use a single feature. The error is evaluated with respect to w_t,

$$\varepsilon_j = \sum_i w_i (h_j(x_i) \neq y_i) \tag{5}$$

   c. Choose the classifier h_t, which has a lowest error ε_t.

   d. Update the weights of samples as

$$w_{t+1,i} = w_{t,i}(\beta^{e_{t,i}}), \tag{6}$$

   where $e_{t,i} = \begin{cases} 1, & h_t(x_i) = y_i \\ 0, & h_t(x_i) \neq y_i \end{cases}$ (7)

   and $\beta_t = \frac{\varepsilon_t}{1-\varepsilon_t}$. (8)

3. The final strong classifier is:

$$h(x) = \begin{cases} 1, & \sum_{t=1}^{T} \alpha_t h_t(x) \geq \sum_{t=1}^{T} \alpha_t \\ 0, & otherwise \end{cases}, \tag{9}$$

where $\alpha_t = \frac{1}{2}\log\frac{1}{\beta_t}$. (10)

5.3. Cascade of Classifiers

Through Adaboost learning algorithm, the classifier at each stage is trained to guarantee the hit rate h and allow a false-positive rate f. Figure 5 shows a cascade of N such classifiers.

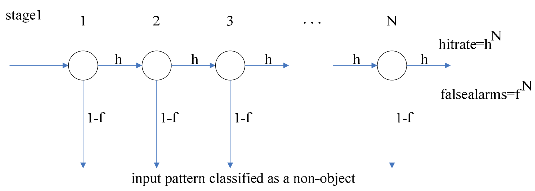

Fig. 5 N-stage cascade structure

For example, if more than 99.9% of object samples are detected and at least 50% (f) of non-object samples are eliminated at each stage, it can be expected that the rate of false alarms is less than and the hit rate is greater than . Obviously, the overall performance is improved by the cascade structure.

6. EXPERIMENTAL RESULTS AND ANALYSIS

In this research, we choose two candidate targets of household objects, i.e., the clock and the computer monitor, which correspond to ellipse and rectangle shapes respectively in the shape detection stage. To apply the Haar-like feature-based recognition to the two household objects, we first collect sample images of clock and computer monitor and use them to train the classifier. The experimental result shows that combining the two stage recognition methods yields better detection rate and less false-positive error than using each single method.

6.1. Experimental Result Using Shape Detection

As mentioned, the robot use a PTZ camera mounted on the platform at the height of 1.1m to capture environmental images. The target object is put on the table (0.73m in height) which is 0.5m~1.5m far from the robot. Within the range of distance, the camera captures 117 test images in our laboratory. The preliminary test result shows that the detection rate is 54% and the false-positive error rate is 5% in detecting the clock object based on ellipse shape; the detection rate is 86% and the false-positive error rate is 6.8% in detecting the monitor object based on rectangle shape, as shown in Figure 6 and Table 1.

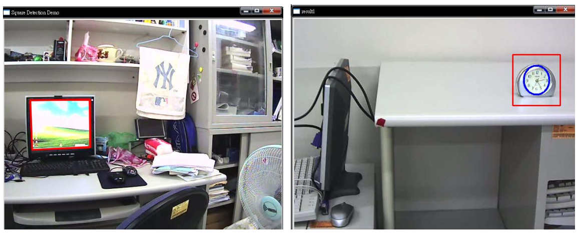

Fig. 6 Experimental result using shape detection

TABLE 1
EXPERIMENTAL RESULT OF HOUSEHOLD OBJECT DETECTION BASED ON SHAPE

	Hits	Missed (false negatives)	False (false positives)	Number of test samples
Clock	54.0%	46.0%	5.0%	117
Computer Monitor	86.0%	13.0%	6.8%	117

6.2. Experimental Results Using Haar-like Features

In this section, the experimental result using Haar-like feature-based recognition is described. First, to collect training samples, the robot captures environmental images with/without target objects in our laboratory, and then it extracts the features of size more than 20x20 pixels but less than 100x100 pixels from captured images. In addition, by applying the geometric transformation to raw object samples, we generate 3000 sample images of size 20x20 pixels to train the classifiers, as shown in Figure 7.

At the same time, 1160 non-object captured images in our laboratory and scenery images searched in Internet are collected to be the background sample images to train the classifiers (shown in Figure 8). Finally, the robot captures 117 images in our laboratory to be testing sample images.

(a) (b)

(c) (d)

Fig. 7 Correct sample images. (a)(b) raw images, (c)(d) extracted images through imageclipper.exe

(a) (b)

Fig. 8 False (background) sample images

Fig. 9 Testing sample images

Training process is performed by haartraining.exe. It is appropriate to set the ratio of correct samples and false samples as 7:3 for training. Thus, the number of correct sample images is 3000 and false sample image is 1160. False alarm is required to be $0.5^{25} \approx 3.0e-08$ and hit rate is required to be $0.986^{25} \approx 0.7$. The size of positive samples is 20x20 pixels. Figure 10 shows that minhitrate must be 0.986 and maxfalsealarm must be 0.5 in each stage of classifier to achieve the expected values when using 25-stage weak classifiers. The weighttrimming is set as 0.95.

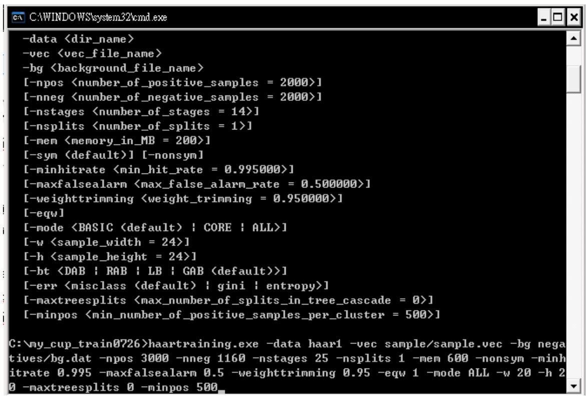

Fig. 10 The parameters used for training through haartraining process

In the training procedures, the most important rectangle features are chosen through AdaBoost in each stage. By matching the rectangle features with the objects in the image, it can determine whether the object is the target. Take computer monitor for example, when the power of the computer is on, the frame of the monitor is darker than the inside. Thus, such feature has better detection rate for computer monitor.

Fig. 11 Rectangle feature chosen from the training procedure

Feature classification of common household objects based on haar-like features is performed to testing samples with household objects through performance.exe. In preliminary test results, the detection rate is 73% and the error rate is 1% in detecting clock based on haar-like features; the detection rate is 52% and the error rate is 11% in detecting computer monitor based on haar-like features (shown in Table 2).

TABLE 2
EXPERIMENTAL RESULT OF HOUSEHOLD OBJECT DETECTION BASED ON HAAR-LIKE FEATURES

	Hits	Missed (false negatives)	False (false positives)	Number of testing samples
Clock	73.0%	27.0%	1.0%	117
Computer monitor	52.0%	47.0%	11.0%	117

6.3. Detection Results Based on Shape and Haar-like Features

17 testing samples with household objects are tested based on shape and haar-like features. Test results show that the detection rate is 75% and the error rate is 0% in detecting clock; the detection rate is 78% and the error rate is 1.7% in detecting computer monitor (shown in Table 3). The result proves that recognition of household object based on the two above methods arises the detection rate and reduce the error.

TABLE 3
EXPERIMENTAL RESULT OF HOUSEHOLD OBJECT DETECTION BASED ON SHAPE AND HAAR-LIKE FEATURES

	Hits	Missed (false negatives)	False (false positives)	Number of testing samples
Clock	75.0%	25.0%	0.0%	117
Computer monitor	78.0%	21.0%	1.7%	117

7. CONCLUSIONS

The research develops a technique to recognize household objects. It can help identify home environment. For domestic robot, it even arise the capability of localization in computer vision.

Thus, the research develops image recognition of based on shape to classify the household objects preliminarily. Then, classification based on haar-like features is also developed to achieve preliminary recognition.

REFERENCES

[1] Wakamaru Communication Robot homepage on Mitsubishi Heavy Industries. [Online] Available: http://www.mhi.co.jp/kobe/wakamaru/english/

[2] Released Press homepage on Toshiba Corporation. [Online] Available: http://www.toshiba.co.jp/about/press/2005_05/pr2001.htm

[3] News on February 5th, 2009 in Physorg.com [Online] Available: http://www.physorg.com/news153079697.html

[4] L. Xu, E. Oja and P. Kultanen, “A new curve detection method: Randomized Hough transform (RHT)”, Pattern Recognition Letters, No. 11, pp. 331-338, 1990.

[5] R. A. McLaughlin, “Randomized Hough transform: better ellipse detection”, IEEE TENCON-Digital Signal Processing Applications, pp.409-414, 1996.

[6] Andrew Schuler, “The randomized Hough transform used for ellipse detection”.

[7] M. Pilu, A. Fitzgibbon, and R. Fisher, “Ellipse-specific direct least-square fitting”, Proc. 7th Int. Conf. on Image Analysis and Processing, Lausanne, Sept. 1996.

[8] Paul Viola and Michael J. Jones. “Rapid Object Detection using a Boosted Cascade of Simple Features,” IEEE CVPR, 2001.

[9] C. Papageorgiou, M. Oren, and T. Poggio. “A general framework for Object Detection,” In International Conference on Computer Vision, 1998.