Vision-based Hand Gesture Recognition for Mobile Service Robot Control

HAILE GEMEDA TAMIRU[1], REN SHU YAN[1, 2], DUAN HAI LONG[1, 2]

1. School of Automation and Electrical Engineering, Tianjin University of Technology and Education, Tianjin 300222, P.R.China

2. Tianjin Key Laboratory of Information Sensing and Intelligent Control, Tianjin, 300222, China

E-mail: gemeda.tami@yahoo.com

Abstract: The vision-based hand gesture is an important problem in the field of human-computer interaction since hand motion and gesture could potentially interact with computers in more natural ways. This paper is focused on control of mobile service robot based on vision-based hand gesture recognition. The database of five hand gesture commands is created and the cross-correlation find the similarity between the captured hand image. This approach contains image pre-processing, segmentation of the region, feature extraction, classification using cross-correlation and gesture recognition. In this method, the user can control the robot by using the gesture of the hand, thereby interacting with the robot system. The commands are generated according to the hand gesture for controlling the robot direction. Experimental results show that this method can get a high recognition rate of hand gesture.

Introduction

In today's age, the robotics industry has been growing up in many new trends to step up the efficiency, accessibility, robustness, accuracy, and computability of the systems. For the main works that are harmful to human and repetitive works that are boring, etc. service robot can be a replacement for human; they still need to be controlled by human itself. They can be wired or wireless, both have the advantage and disadvantage associated with them. Beyond controlling the system through physical devices, a recent method of using gesture to control the robot was very popular. The main target of using gesture is that it delivers a more natural way of controlling. This mainly includes image processing and machine learning for the system. Beyond this, for gesture control also requires some kind of hardware for interfacing with a system. There are some systems that have been upgraded in the same field using many techniques. From different gesture recognition system, this paper is based on hand gesture recognition.

Hand gesture recognition is an active research area where we have a large scope for improvements and inventions. Gesture is physical movements which have a meaningful information so it is nonverbal communication in a natural way. In the past two decades gesture recognition has become an influencing term. For tracking, extracting, and recognizing hand gestures many gesture recognition techniques have been developed. Each one of them has its own advantages and disadvantages as well. The First is wired technology, which is interfaced with the computer system. In wired technology, the user can not freely move because they are limited by the length of wires to cover the distance which connects to the computer system via wires. One of the wired technologies is instrumented gloves (electronics, gloves and a data glove) [1]. It is used to provide the information related to hand location, orientation which contains some sensors. The data gloves are very expensive to utilize in a broad range of

application, but provide results with high accuracy. Optical markers are another choice to replace data gloves. An optical marker provides the information about the exact location of the hand or tips/knuckles of fingers by projecting Infra-red light and reflecting this light on the screen. These systems also require very complex configuration and have a good result.

The second advanced techniques which are introduced to vision-based techniques [2] require processing of images features like texture, color, etc. The result of hand gesture recognition may vary and would be different as skin color changes from person to person and from one country to the other if color texture feature of the image are used. And also color, texture gets modified under different illumination condition. Shape-based features for hand gesture recognition are another alternative for the same purpose, and we reached to employ different. This is a universal truth that under normal condition every person poses almost the same hand shape with one thumb and four fingers. The hand shape-based gesture recognition frame rate achieved is comparable to most of the system in existence, but the number of different gestures recognized and the recognition accuracy are amongst the best found.

For controlling mobile service robot based on hand gesture recognition many systems exist. Some hand gesture recognition systems involve, adaptive color segmentation [3], labeling, and a hand finding and with blocking morphological filtering, and then gestures action is found by skeleton sing and template matching. For hand gesture inputs due to template matching, this does not provide dynamicity. Another system uses machine interfacing device to provide real-time gestures to the robot [4]. An analog flex sensor is used to use in the hand glove to measure the finger blending [5], also hand position and orientation are measured by ultrasonic for gesture recognition [6]. And in another approach, gestures are recognized using Microsoft Xbox kinetic (c) [7]. Kinetic gathers the color and depth information using an RGB and Infra-Red camera respectively. This system thought is not very cost effective.

The main goal of this paper is to provide a real-time (reliable) and more natural method for the user to control the robot. A vision-based hand gesture recognition approach is proposed for controlling the direction of the wireless robot in the environment through five hand gesture commands. Hand gestures are snapshot on real-time to generate commands through a good quality inbuilt camera (webcam) or USB 2.0 webcam. Then, the robot is moving in all possible directions in the environment using five possible types of commands which are Right, left, forward, backward, and stop. The contribution of this paper is using a real time hand gesture recognition with an accuracy rate of 95.2%, and using a simple and clear algorithm.

The remainder of this parer is organized as follows. The database is described in Section 2. Section 3 conducts system implementation. The algorithm is formulated in Section 4. Validation experiments are presented in Section V. The paper concludes in Section 6.

Database description

Our approach for hand gesture recognition is based on static mode; Our first problem is to gather a good quality of data since our classifier will classify character according to it only. We had created our own database for each action and gesture which include 50 images, i.e. 10 images for each gesture, as shown in Fig. 1.
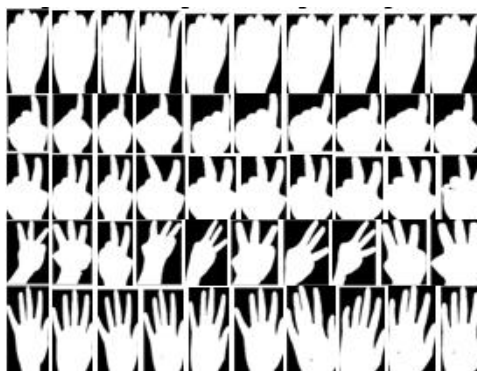
Fig. 1 Database of five hands

System implementation

We describe the system implementation in detail in this section. The block diagram of the system is shown in Fig.2.
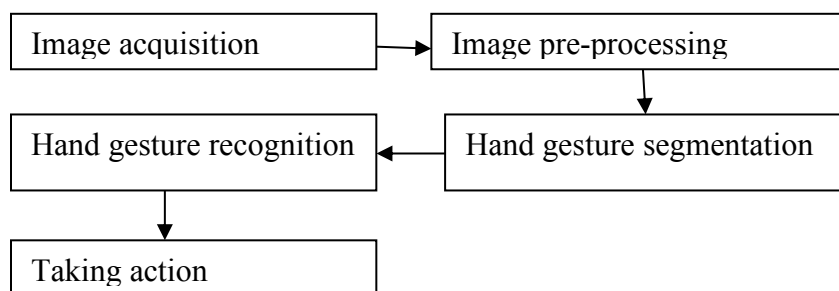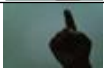

Fig. 2 Block diagram of the system

Image acquisition: Image acquisition block enables to acquire images and video from webcam camera into MATLAB and we can detect hardware directly. In acquiring image it results in the background and synchronizes various multimodal devices.

Image pre-processing: The process of enhancing the quality of images is Histogram equalization. This can be achieved by transforming the values in the color map. The specified histogram is being matched with output image. For best accuracy, noise must be removed from the image. In this paper, for removing noise median filtering can be used.

Table 1 hand gesture and the expected result

No.	Hand gesture	Action
Zero		Stop
One		Forward direction
Two		Left direction
Three		Right direction
Five		Backward direction

Hand gesture segmentation: Segmentation is a process in which we convert an RGB image to a grayscale image, and grayscale image into a binary image (black and white). This is to be done because we can get only two objects only in our image which is black and white. Otsu algorithm is used to convert the image into binary; white represents our hand and black with the background. To get the best result we have to select a best threshold value according to this value and segmentation can be done. The segmentation process mainly depends on the type of hand image on which we have to do the processing and Otsu algorithms had been tested with our hand gestures data work efficiently. It can choose a threshold value automatically and do segmentation because it is a nonparametric and an unsupervised method of segmentation**.** After applying this algorithm the segmented images need further processing is on those images to remove errors and unwanted data. There are still some blacks (0s) on hand parts and some whites (1s) on background parts. To remove this noise from those segmented images we need to apply morphological filtering techniques. It is necessary to remove these errors as they can create a problem in recognition of hand gestures and reduce the system efficiency. So morphological filtering is necessary to apply to segmented images and then we get a better contour of a gesture smooth closed. Erosion, Opening, Dilation, and Closing are the basic operators that work in morphological filtering. A sample of the pre-processing result is shown in Fig. 3, and the experiments are performed in MATLAB. After the pre-processing, we get a smooth and better hand gesture which can result in a better efficiency.

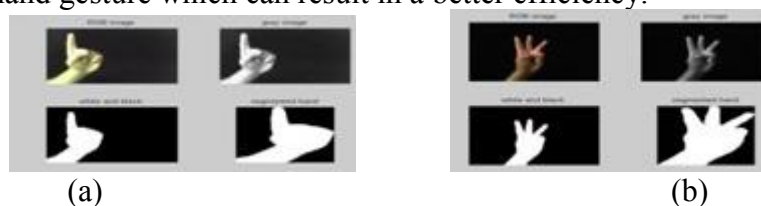(a)   (b)

Fig. 3 the sample of hand segmentation

Hand gesture recognition: Gesture recognition is basically a term used for determining human gestures with mathematical algorithms. The gesture can originate from body movement or state, but normally originates from the hand. It bridges the gap between machine and humans with primitive text, also called as a GUI (graphical user interfaces). It also has demerits with input to the keyboard as well as a mouse. Gesture recognition provides a way for humans to interact with the machine and communicate naturally.

The cross - correlation coefficient is calculated to recognize hand gestures. The Cross-correlation 2D object computes 2-D cross-correlation of two input matrices. It has its maximum value when the two matrices are aligned so that they are shaped as similarly as possible. Now we have to extract feature for gesture recognition. For matching the hand gesture, we used Cross-correlation Coefficient. The cross-correlation coefficient is defined by Equation (1).

$$\gamma(x,y) = \frac{\sum_s \sum_t \delta I(x,s,y+t)\delta\,\delta T(s,t)}{\sum_s \sum_t \delta^2 I(x+s,y+t)\delta^2\,\delta T(s,t)} \dots\dots\dots \text{(1)}$$

Where $\delta I(x+s, y+t) = I(x+s, y+t) - I'(x, y)$

$$\delta_T(s, t) = T(s, t) - T',$$

$$s \in \{1,2,3,...., p\},$$

$$t \in \{1,2,3,...., q\},$$

$$x \in \{1,2,3,...., m-p+1\},$$

$$y \in \{1,2,3,...., n-q+1\},$$

$$I'(x, y) = \frac{1}{pq} \sum_s \sum_s I(x+s, y+t)$$

The value of γ ranges from -1 to +1 and corresponds completely not matched and completely matched respectively. For matching the template T slides over I and γ is calculated for each coordinate (x, y). After calculation, the point which gets maximum γ is referred to as the match point.

1. Algorithm

The procedure of this algorithm is described as follows.

1. Start
2. Image acquisition.
3. Image pre-processing.
   - ✓ RGB to gray
   - ✓ Gray to black and white
   - ✓ Noise reduction
   - ✓ Histogram adjustment
   - ✓ Trace the region boundaries
   - ✓ Create a mask
   - ✓ Morphological processing
4. Thresholding based hand segmentation.
   - ✓ Otsu algorithm
5. Hand Gesture recognition.
   - ✓ Using the cross - correlation coefficient.
6. Controlling based on the recognized gesture.
   - ✓ Left direction
   - ✓ Right direction
   - ✓ Forward direction

    ✓ Backward direction

    ✓ Stop

7.   Repeating step 2 to 6 till exit.

8.   Stop.

Experimental Result

In this section, first 50 hand gesture images is stored in the database as shown fig. 1, second we train our hand gesture as shown in fig 4.a, 5.a, 6.a, and 7.a. And the last part is checking the similarity of the trained data and the database by the cross-correlation which is called the gesture recognition result as shown in fig 4.b, 5.b, 6.b, and 7.b. We evaluate the performance of the proposed method using MATLAB 2015 environment on an Intel Core i7 with 4GB of RAM running.

Fig 4.a trained the hand gesture   Fig 4.b gesture recognition.

Fig 4.a and 4.b show when the camera capture the background, capture gesture image with a mask, segment, train, and recognize it as left direction.

Fig 5.a train the hand gesture   Fig 5.b gesture recognition.

Fig 5.a and 5.b show when the camera capture the background, capture gesture image with a mask, segment, train, and recognize it the right direction

Fig 6. A trained the hand gesture   Fig 6. b gesture recognition.

Fig 6.a and 6.b show when the camera capture the background, capture gesture image with a mask, segment, train, and recognize it as a stop.

Fig 7.a trained the hand gesture   Fig 7.b gesture recognition.

Fig 7.a and 7.b show when the camera capture the background, capture gesture image with a mask, segment, train, and recognize it as the backward direction.

In this work, for experimental analysis, we had applied the above method to our database of the hand gesture 500 times of hand image, i.e. 100 hand images per gesture and as able to recognize all of the hand gesture. Table 2 shows the accuracy rate for each of the hand gesture

Table 2 the accuracy rate

Hand gesture	Input image	Recognized image	Accuracy rate (%)
	100	97	97
	100	95	95
	100	94	94
	100	91	91
	100	99	99
Total	500	476	95.2

The results suggest that our method achieves excellent performance for all the cases.

Conclusion

In this paper, a vision-based hand gesture recognition approach for controlling mobile service robot is proposed. We interface the camera with MATLAB for real-time image acquisition. First, we captured the hand gesture by the camera and then we have extracted the hand gesture from the main image using the segmentation method which is known as Otsu algorithm method. Here morphological filtering techniques are applied to remove the noise. A cross-correlation coefficient is used in feature extraction stage to recognize the real hand gesture. This algorithm identifies hand gesture by measuring the similarity of the images (the trained hand gesture, and the captured hand gesture in a given mask.) using cross-correlation coefficient value of 50 database hand gesture image and snapshot gesture. Accordingly, the gesture recognition is known, and identifies the direction. The experimental results show the validity of the proposed approach. Our future work is focused on implementation of the algorithm on mobile service robots and exploring of another type of gesture with a different algorithm.

Reference

[1] Haitham Hasan, Sameem Abdul-Kareem, "Human–computer interaction using vision-based hand gesture Recognition systems: a survey", Neural Comput & Applic, 2013.

[2] Fahn, C.S, Sun, H, "Development of a fingertip glove equipped with magnetic tracking sensors", Sensors 2015, Vol. 10, pp. 1119–1140,2015.[3] Chao Hy Xiang Wang, Mrinal K. Mandal, Max Meng, and Donglin Li, "Efficient Face and Gesture Recognition Techniques for Robot Control", CCECE, 1757-1762, 2003.

[4] Asanterabi Malima, Erol Ozgur, and Mujdat Cetin, "A Fast Algorithm for Vision-Based Hand Gesture Recognition for Robot Control", IEEE International Conference on Computer Vision, 2006.

[5] Prof. Chitte p.p., Miss: Khemnar S.B., Miss: Kanawade A.A., Miss: Wakale S.B., " A hand gesture based wheelchair for a physically handicapped person with emergency alert system", IRJET Volume: 03 Issue: 04 | Apr-2016.

[6] Jagdish Lal Raheja, Radhey Shyam, Umesh Kumar and P Bhanu Prasad, "Real-Time Robotic Hand Control use Hand Gestures", Second International Conference on Machine Learning and Computing, 2010.

[7] Gesture Controlled Robot uses Kinect http://www.e-yantra.org/home/projects-wiki/item/180-gesture controlled-robot-using-Firebird-and-Kinect.

[8] Fahn, C.S, Sun, H, "Development of a fingertip glove equipped with magnetic tracking sensors", Sensors 2015, Vol. 10, pp. 1119–1140,2015.