Hand gesture recognition for human-computer interaction

[1]Mayur Yeshi, [2]Pradeep Kale, [3]Bhushan Yeshi, [4]Vinod Sonawane

[1]Dept. of Mechanical Engg. , ICOER, Savitribai Phule Pune University, Pune, India.
[2]Dept. of Mechanical Engg. , School of Engineering & Technology, Madda walabu University, Ethiopia.
[3]Associate Engineer, Tech Mahindra, Pune, India.
[4]Safety Officer at Siddhivinayak Aesthetics (PVT.) Limited, Chakan, Pune, India.

***ABSTRACT*- This paper focus on advance study of Gesture control based robot. The first part of the paper provides an overview of the current state of the art regarding the recognition of hand gestures as these are observed and recorded by typical video cameras. We derive a set of motion features based on smoothed optical flow estimates. A usercentric representation of these features is obtained using face detection, and an efficient classifier is learned to discriminate between gestures. A number of hand gesture recognition technologies and applications for Human Vehicle Interaction (HVI) are also discussed including a summary of current automotive hand gesture recognition research.**

***Index Terms*- Gesture control, hand gesture, Human vehicle interaction, skin filtering.**

I. INTRODUCTION

Hand gesture recognition offer potential safety benefits for various types of secondary controls. Face, head and body gesture recognition technologies may also offer some safety benefits [1]. In Skin filtering the RGB image is converted to HSV image because this model is more sensitive to changes in lighting condition [2]. Human gesture recognition in image sequences has many applications including human-computer interaction, surveillance, and video games [3].The somatosensory interaction is one of the most user-friendly interactive interfaces for controlling objects.Motivated by the idea of a Wiimote, we try to implement interface which allows a user to navigate a car-robot in asomatosensory interactive way. An easy way is to directly use a Wiimote to control a robot; however, the price animate is not very low and Wiimote's size is not very small either. Therefore, the interface developed by us adopts a smallsized accelerometer module instead of the traditional Wiimote [4]. Recently, there have been many different hand gesture recognition systems, such as vision-based trajectory recognition systems and inertial-based trajectory recognition systems. No matter cameras or accelerators are used in the hand gesture systems; the core module is a hand gesture recognition algorithm [5].

II. LITERATURE REVIEW

A literature review of current research investigating the use ofhand gestures for vehicle secondary controls has been carriedout and is briefly summarized in the following section. Thissummary presents the different technologies and techniques used by different researchers. Previous research does notfocus on understanding driver behavior or the limitations ofhand gestures.The literature review and the resulting analysis led to the proposed classification of the research [1].There are many approaches that were followed by different researchers like vision based, data glove based, Artificial Neural Network, Fuzzy Logic, Genetic Algorithm, Hidden Markov Model, Support VectorMachines etc. Some of the previous works are given below. Many researchers usedVision based approaches for identifying hand gestures and found out the skin colored region from theinput image captured and then this image with desired handregion was intensity normalized and histogram was foundout for the same. Feature extraction step was performedusing Hit-Miss Transform and the gesture was recognized using Hidden Markov Model [2-4].The prime aim of the design is that the robot and platform starts the movement as soon as the operator makes a gesture or posture or any motion. The Robotic arm is synchronized with the gestures (hand postures) of the operator and the platform part is synchronized with the gestures (leg postures) of the operator [5-6].This paper describes a gesture interface for the control of a mobile robot equipped with a manipulator. The interface uses a camera to track a person and recognizegestures involving arm motion. A fast, adaptive tracking algorithm enables the robot to track and follow a person reliably through environments with changinglighting conditions. Two alternative methods for gesture recognition are compared:a template based approach and a neural network approach. Both are combined withthe Viterbi algorithm for the recognition of gestures defined through arm motion (inaddition to static arm poses). Results are reported in the context of an interactiveclean-up task, where a person guides the robot to specific locations that need to be cleaned and instructs the robot to pick up trash [7-9].Mobile robots have the capability to move around in their environment and are not fixed to one physical location.The movement can be achieved using legs, wheels or other different mechanism. They have the advantage of consuming lessenergy and move faster than other type of locomotion mechanisms. Hand gesture recognition system plays an important rolein the human-robot interactions, due to the fact that hand gestures are a natural and powerful way of communication, andcan be used for the remote control of robots. Two approaches are commonly used to interpret gestures for human robotinteraction, gloved-based approach and vision-based method. Using gloves, it requires wearing of cumbersome contactdevices and generally carrying a load of cables that connect the device to a computer [10-12].

III. Computer Vision Techniques for Hand Gesture Recognition

A) Gesture

Issuing commands to robot through only computer vision without sounds or other media is similar to conducting a marching band by way of visible gestures. In order to achieve real-time operations, our system requires simpler body language which is easy to recognize and differentiate from each other. Our system is mainly to recognize the dynamic hand gesture from continuous hand motion in real-time, and implement on interaction between human and robot. There are very many kinds gesture can be represented by the hand motion. In this system, we describe four types of directive gesture to one hand, which is moving upward, moving downward, and moving leftward and moving rightward separately, for the basic conducting gesture. Thus, if we add one or both hands into gesture invoking, we will have at most twenty-four kinds of meaningful gesture by the permutation combination of both hands. Here, we use a 2-D table to represent the all combination of gesture from both hands and classify each combination into a class which is named gesture's ID individually. By the way is easy to represent every gestures and it is convenient to add new hand gestures [2]. Most of the complete hand interactive systems can be considered to be comprisedof three layers: detection, tracking and recognition. The detection layer isresponsible for defining and extracting visual features that can be attributed tothe presence of hands in the field of view of the camera(s). The tracking layer isresponsible for performing temporal data association between successive imageframes, so that, at each moment in time, the system may be aware of \whatis where". Moreover, in model-based methods, tracking also provides a way to maintain estimates of model parameters, variables and features that are notdirectly observable at a certain moment in time. Last, the recognition layer isresponsible for grouping the spatiotemporal data extracted in the previous layersand assigning the resulting groups with labels associated to particular classesof gestures. In this section, research on these three identified sub-problems ofvision-based gesture recognition is reviewed [6].

B) Recognition

The overall goal of hand gesture recognition is the interpretation of the semantics that the hand(s) location, posture, or gesture conveys. Basically, there have been two types of interaction in which hands are employed in the user's communication with a computer. The first is control applications such as drawing,where the user sketches a curve while the computer renders this curve on a 2DMethods that relate to hand-driven control focus on the detection and tracking of some featureand can be handled with the information extracted through the tracking of these features. The second type of interaction involvesthe recognition of hand postures, or signs, and gestures. Naturally, the vocabulary of signs or gestures is largely application dependent. Typically, the largerthe vocabulary is, the hardest the recognition task becomes. Two early systemsindicate the difference between recognition [BMM97] and control [MM95]. The first recognizes 25 postures from the International Hand Alphabet, while the second was used to support interaction in a virtual workspace [8].

C) Hand Gesture Interfaces

Gesture interfaces can range from those that recognize a fewsymbolic gestures to those that implement fully fledged signlanguage interpretation. Gesture interfaces may also recognize static hand poses, or dynamic hand motion, or a combinationof both. In all cases each gesture should have an unambiguoussemantic meaning associated with it that can be used in theinterface [1].However, this Paper will address only one specific use of the term "gesture" that is, hand gestures that are considerednatural or co-occur with spoken language. This narrow focus is because the author fully agrees with the views expressed byCassel, who states that she does not believe that everydayhumans have a natural affinity for a learned "gestural language". Natural hand gestures are primarily found inassociation with spoken language, 90% of gestures are foundin the context of speech according to McNeil [9]. Thus, ifthe goal is to get away from learned, pre-defined interactiontechniques and create naturaland safe interfaces free ofvisual demand for normal human drivers, then the focusshould be on the type of gestures that come naturally tonormal humans.Therefore, this Paper is focused on discussing the use ofnatural, dynamic non-contact hand gestures only and although safety is the primary motivation for this research, other automotive applications will also be mentioned [2].

D) Hand Gesture based Secondary Controls

1. Application Domains: A) Primary Task controls in vehicle Device, e.g.Radio, CD, Phone, Navigation, etc. B) In-vehicle control type, e.g. Push button switch, Rotary selector, Slider, Cursor control, touch panel, etc. C) Selective theme mapping, e.g. Lighting, Closures, Context sensitive, Visualmanual tasks, etc.

2. Hand Gesture System Design Techniques: A) Multimodal B) Unimodal C) ContactD) Non-Contact E) Dynamic F) Static G) One-handed H) Gesture location upper

I) Gesture location lower J) Driver Visual reminder K) Driver Feedback.

3. Hand Gesture Only Interface Type: A) Natural B) Symbolic C) Sign Language.

4. Hand Gesture System Design Technologies: A) Intrusive B) Non-Intrusive C) Device-based D) Vision-based E) Sensor-based.

IV. Sensor-based Technologies

In order to find a solution that can be used today and one that overcomes some of the limitations of vision-based systems [10]. Limited research has been carried out using a sensor-based approach. Lasers have been used for gesture musical applications; however no automotive use of lasers for hand gesture applications has been found [11]. Capacitive andinfra-red techniques have been referred to in some literature, but no evidence or publications of working handgesture systems has been identified. Electric Field Sensing (EFS) was initially pioneered byMassachusetts Institute of Technology (MIT), and can detectthe presence of a human hand on or near to a conductiveobject. EFS are not affected by dynamic backgrounds orvariable lighting conditions and has very fast response times [12]. However, the system limitations of EFS for in-vehicleapplications have been investigated by this research viaexperimentation and EFS was found to be sensitive to the userbeing earthed, thickness of clothing worn, water, contact withother person within the vehicle and there wascomputational difficulty in locating a hand in 3D. Afterfurther research, all the above technical difficulties wereeventually resolved and EFS was initially used as the gesture technology for the early stages of this research. As researchprogressed a range of simpler and less capable sensor-basedtechnologies were also investigated for specific applications [3].

A) Face Detection

Face detection is used to create a normalized, usercentricview of the user. The image is scaled based onthe radius of the detected face, and is then cropped andcentered based on the position of the face.

B) Block diagram

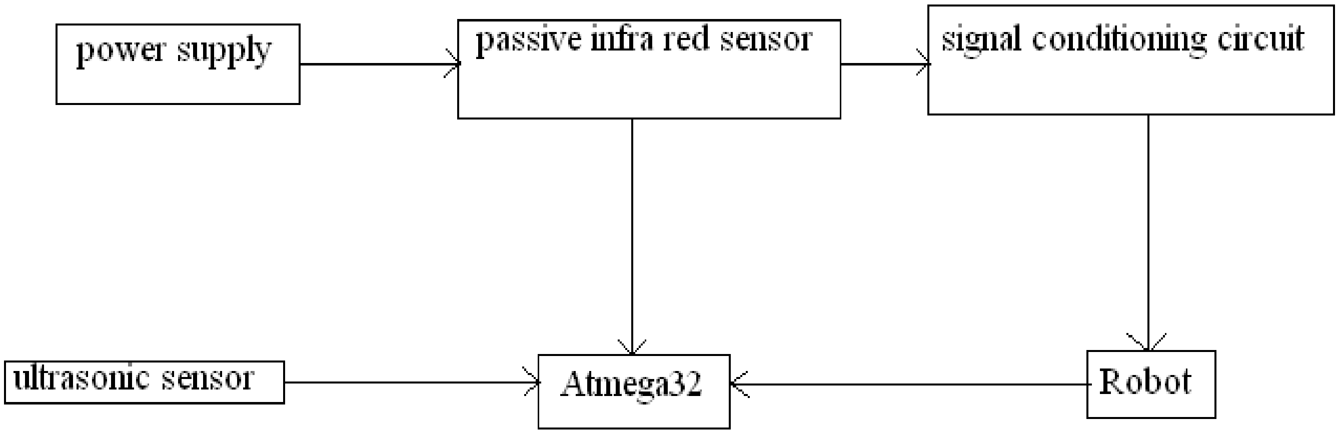

Fig.1 Block diagram of gesture control based robot

C) Hardware Component

Fig. 2 illustrates the overall view of the system. The system contains of a camera as the vision sensor connected to a laptop for processing image/video. The laptop is connected to XBee wireless transmitter module which in turn sends data to another XBee receiving platform which receives the data and in turn sends data to the main controller containing a PIC chip, which is interfaced to the robot motors through a motor driver circuit.

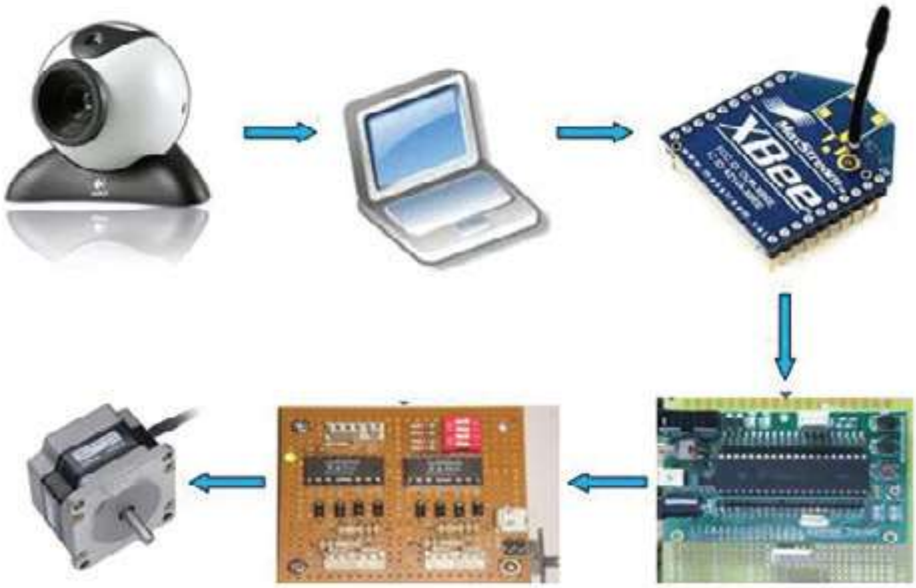

Fig 2 Hardware component

D) Hand Control Using Hand Gesture

Technique to acquire hand gestures and to control robotic system using hand gestures is:

A) Sensor acquisition to get hand gestures.
B) Extracting hand gesture area from captured frame.
C) Generation of instructions corresponding to matched gesture, for specified robotic action.

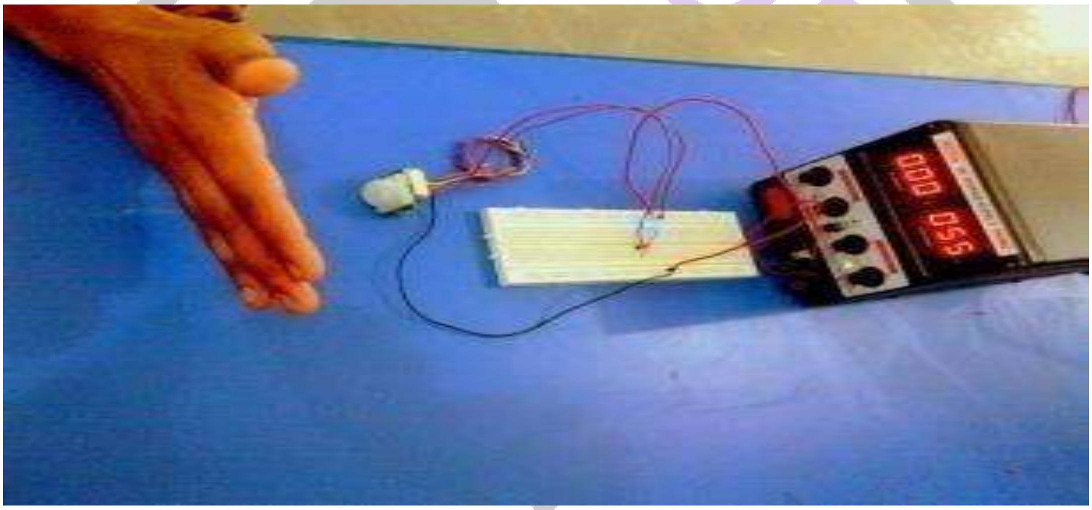

Fig.3 Hand control using hand gesture

E) Image filtering

Skin color segmentation can effectively segment the skin color pixels, but it gives noisy holes in the image due to nails and other hand accessories like ring. So, morphological image filtering is used to filter out the noisy pixels. The morphological operation performed on the image is described. A B = (A*B) B

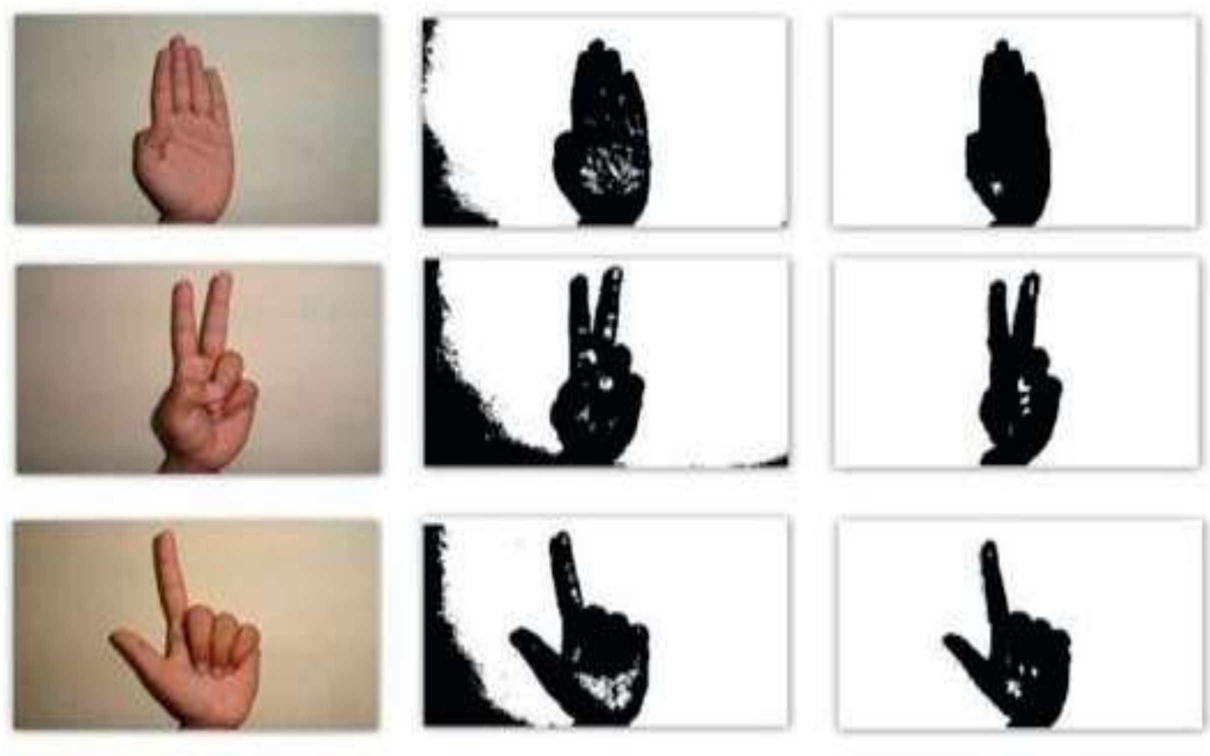

Fig.4 Image filtering

F) Color Detection

The custom color is predefined to be red, where only the red color channel is used and the other channels contributes forany deviation of red color. The RGB Filter filters out the red pixels from the object and eliminate for any other colorswhere

$$R = (R-G) + (R - B),$$
$$G = 0, B = 0$$

Thresholding property is also available to eliminate for pixels with dark or white color not containing enough red color. Ahysteresis level is applied to the object for eliminating any red noises produced by lighting or variations of intensity. In the red hand glove is segmented from the background for further processing.

V. CONCLUSION

In this paper, we studied a hand-gesture-based interfacefor navigating a robot. A user can control a robotdirectly by using his or her hand trajectories. In the future, wewill directly use a mobile phone with an accelerometer tocontrol a robot. We also want to add more hand gestures(such as the curve and slash) into the interface to control in a more natural and effectively way.The hand gestures will be standardized in the future so that the command can be implemented throughout the world without worrying about cultural and language differences. In other words, it is our goal to develop commands based on standard hand gestures for universal applications.

REFERENCES

[1] Carl A. Pickering, "A Research Study of Hand Gesture Recognition Technologies andApplications for Human Vehicle Interaction", Automotive Electronics, 2007 3rd Institution of Engineering and Technology Conference, June 2007, pp. 1-15.

[2] JoyeetaSingha and Karen Das; "Hand Gesture Recognition Based on Karhunen-Loeve Transform", IEEE, Mobile & Embedded Technology International Conference 2013, pp.366

[3] Mark Bayazit, Alex Couture-Beil, Greg;"Real-time Motion-based Gesture Recognition using the GPU", in Proc. of the IAPR conf. on Machine Vision Applications, 2009, pp.9-12.

[4] X. Zabulisy, H. Baltzakisy, "Vision-based Hand Gesture Recognition for Human-Computer Interaction", IEEE, In: The Universal Access Handbook. LEA, 2009.

[5] Dr. Gowri Shankar Rao, Dr. D. Bhattacharya, Ajit Pandey, Aparna Tiwari; “Dual sensor based gesture robot control using minimal hardware system”, International Journal of Scientific and Research Publications, Volume 3, Issue 5, May 2013, ISSN 2250-3153.

[6] Love Aggarwal B.Tech (ECE), GGSIPU; “Design and Implementation of a Wireless Gesture Controlled Robotic Arm with Vision”, International Journal of Computer Applications, Volume 79 – No 13, October 2013, pp.39-43.

[7] Ahmad 'AthifMohdFaudzi, “Real-time Hand Gestures System for Mobile Robots Control”, IRIS 2012, Procedia Engineering 41, pp. 798-804.

[8] Stefan Waldherr Computer Science Department Carnegie Mellon University Pittsburgh, PA, USA ; “A Gesture Based Interface for Human-Robot Interaction”, Autonomous Robots September 2000, Volume 9, Issue 2, pp. 151–173.

[9] Ming-Shaung Chang; “Establishing a natural HRI System for Mobile Robot Through Human Hand Gestures”,IFAC Proceedings Volumes Volume 42, Issue 16, 2009, pp. 723-728.

[10] Chang-Yi Kaoa* and Chin-ShyurngFahn; “A Human-Machine Interaction Technique: Hand Gesture Recognition Based on Hidden Markov Models with Trajectory of Hand Motion”, CEIS 2011 Procedia Engineering Volume 15, 2011, pp. 3739-3743.

[11] Amit Gupta, Vijay Kumar Sehrawat, Mamta Khosla; “FPGA Based Real Time Human Hand Gesture Recognition System”, Amit Gupta et al. / Procedia Technology 6 (2012), pp. 98 – 107.

[12] Mu-Chun Su National Central University; “A hand-gesture-based control interface for a car-robot”, The 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems October, 2010, pp. 18-22.