IJCSC International Journal of Computer Science and Communication Vol. 2, No. 1, January-June 2011, pp. 223-228

MULTIMODAL INTERFACE FOR DISABLED PERSONS

S.A. Chhabria and R.V. Dharaskar

Department of Computer Science and Engineering, G.H. Raisoni College of Engineering, Nagpur.
E-mail: sharda_chhabria@yahoo.co.in

ABSTRACT

Assistive robots have the potential to provide disabled people with effective ways to alleviate the impact of their limitations, by compensating for their specific impairments. In particular, robotic wheelchairs may help in maneuvering a wheelchair and planning motion. The eye tracking systems are the most widely used and efficient way to control the systems used by the people with motor disabilities. The eye controlled wheelchair offers yet another alternative to persons who cannot use joysticks. Gaze-based human-computer interaction (HCI) enables users to operate computers by means of eye movement.

In a multimodal conversation, the way users communicate with a system depends on the available interaction channels and the situated context (e.g., conversation focus, visual feedback). A correct interpretation can only be attained by simultaneously considering these constraints.

This paper presents a technique of how Eye, Hand Gestures or voice command can be used to control the movement of wheelchair. For Hand gestures we present a scheme which reduces the size of the database which is used to store different postures of human beings which are used by robot as commands. The picture frame may divide into different scan lines and pixel color value under these scan lines are examined to guess the particular posture of user. The robot may use this as command and act accordingly. For interfacing through eye a new algorithm for tracking the movement of eye towards left or right is also proposed. The eye blinking feature can also be used by the algorithm to control the starting and stopping of wheelchair.

This paper presents a technique for developing user friendly multimodal interface using eye, or hand gestures to control the movement of wheelchair.

Keywords: Hand Gestures, Eye Movement, Human-computer Interaction (HCI)

1. INTRODUCTION

Human-robot symbiotic systems have been studied extensively in recent years, considering that robots will play an important role in the future welfare society [Ueno, 2001]. The use of intelligent robots encourages the view of the machine as a partner in communication rather than as a tool [1]. In the near future, robots will interact closely with a group of humans in their everyday environment in the field of entertainment, recreation, health-care, nursing, etc. In human-human interaction, multiple communication modals such as speech, gestures and body movements are frequently used. The standard input methods, such as text input via the keyboard and pointer/location information from a mouse, do not provide a natural, intuitive interaction between models for natural and intuitive communication between humans and robots. Furthermore, for intuitive gesture-based interaction between human and robot, the robot should understand the meaning of gesture with respect to society and culture. The ability to understand hand gestures will improve the naturalness and efficiency of human interaction with robot, and allow the user to communicate in complex tasks without using tedious sets of detailed instructions. This interactive system uses robot eye's cameras or cameras to identify humans and recognize their gestures based on face and hand poses.

The population of people with disabilities has risen markedly during the past century. As the data come from the National Health Interview Survey (NHIS), two distinct trends have contributed to the increasing overall prevalence of disability: a gradual rise, due largely to demographic shifts associated with an aging population, as well as a rapid increase that is due to health impairments and accidents [2].

Many individuals have problems to use a conventional wheelchair. A recent clinical survey [3] indicated that 9%-10% of patients who received power wheelchair training found it extremely difficult or impossible to use it for their activities of daily living, and 40% of patients found the steering and maneuvering tasks difficult or impossible. These people, suffering from motor deficits, disorientation, amnesia, or cognitive deficits, are dependent upon others to push them, so often feel powerless and out of control.

Assistive robots [4] have the potential to provide these people with effective ways to alleviate the impact of their limitations, by compensating for their specific impairments. In particular, robotic wheelchairs [5] may help in maneuvering a wheelchair and planning motion. Recently, research of assistant robots is also emerging field of robotic applications.

The remainder of the paper is structured as follows. Section 2 describes literature survey. Section 3 discusses proposed plan for system and, finally, Section 4 presents concluding remarks and future directions.

2. LITERATURE SURVEY

Multimodal interpretation is a process of identifying semantic meaning from user inputs.

For the system where speech is the main mode of communication and accompanied by deictic gestures, **a semantic fusion** approach is widely adopted for multimodal interpretation.

Salience driven approach for input interpretation in multimodal conversational systems [6] takes advantage of rich information from multiple modalities. Information from deictic gestures is used to identify a part of the physical world that is salient at a given point of communication. This salient part of the physical world is then used to prime language models for spoken language understanding. This approach reduces word error rate and improves concept identification from spoken utterances. Although currently this approach has only investigated the use of gesture information in salience modeling, the salience driven approach can be extended to include other modalities (e.g., eye gaze) and information (e.g., conversation context). Future work will specifically investigate how to combine information from multiple sources in salience modeling and how to apply the salience models in different early stages of processing [6].

In Shoushun Chen paper [7] the architecture of the proposed system is illustrated in Fig. 1. It includes an image sensor working at temporal difference mode, a hierarchical edge feature extraction unit and a classifier with a set of library postures.

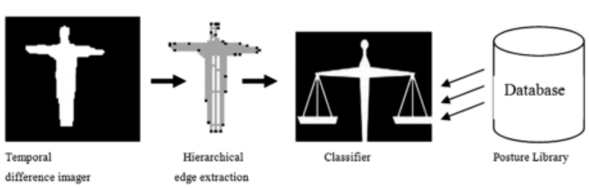

Fig. 1: Block Diagram of the System

The temporal difference image sensor compares two continuous image frames and only outputs addresses of those pixels with illumination variance larger than certain threshold. If the scene illumination and object reflectance are constant then the changes in scene reflectance only result from object or viewer movement. Therefore the background information is naturally filtered since the received pixels only come from the active object of interest. This shows great computational efficiency as compared to conventional image sensors used in other systems. With the address of the events, an edge feature extraction unit will reorganize the contour of the objects into vectorial line segments. The extracted line segments are fed to a modified Hausdorff distance scheme to measure the similarity of the input line segments with those of a set of library objects.

This classifier is able to perform size and position invariance recognition from object or viewer movement. Therefore the background information is naturally filtered since the received pixels only come from the active object of interest. This shows great computational efficiency as compared to conventional image sensors used in other systems. With the address of the events, an edge feature extraction unit will reorganize the contour of the objects into vectorial line segments. The extracted line segments are fed to posture library.

Based on the physical disability of the user, different wheelchair control methods are used. The wheelchair control methods are dependent upon the way the commands are interpreted form the user. The different technologies for controlling the wheelchair are Electrooculographic Potential (EOG), Oral Motion Controlled Intelligent Wheelchairs, and Control of Wheelchair by Head Movements, Control of Wheelchair by Hand Gestures.

The Literature Survey had revealed with the widespread use of computers in modern society, traditional human-computer interaction (HCI) technologies based on mouse and keyboard show their increasing limitations. Thus, research on multimodal HCI is becoming more and more important in real life. Design of intelligent multimodal user interface is one of the important research areas of HCI, has spawned more and more interest in HCI society. So we may try to think of going with further research using suitable techniques.

3. PROPOSED PLAN

This paper presents a technique of how Eye or Hand Gestures can be used to control the movement of wheelchair. For Hand gestures we present a scheme which reduces the size of the database which is used to store different postures of human beings which are used by robot as commands. The picture frame may be divided into different scan lines and pixel color value under these scan lines are examined to guess the particular posture of user. The robot may use this as command and act accordingly.

For interfacing through eye a new algorithm for tracking the movement of eye towards left or right is also proposed. The eye blinking feature can also be used by the algorithm to control the starting and stopping of wheelchair.

This paper presents a technique for developing user friendly multimodal interface using eye, hand gestures or voice command to control the movement of wheelchair.

3.1 Interface through Eye

3.1.1 Eye Movement Tracking

The eye motion tracking hardware includes a USB web camera which is mounted on a cap worn by the user. This camera is adjusted so that it lies in front of one of the eye of user. The camera has inbuilt light source, so that it can capture bright images if darkness appears under the cap.

The drivers of the camera are installed in a PC to which the camera is plugged in. The software module for image processing works on three different modules: video capturing, frame extraction and pixel color detection. The image processing program performs these three steps based on the coherence algorithm explained below.

3.1.2 Coherence Algorithm

The coherence algorithm works for detecting the motion of eye. This algorithm operates on the frames extracted from the video of the eye. From the frame, the algorithm extracts the pixels which lie on the vertical edges of the rectangular area selected by the user. These pixels are then processed to determine the RGB values. When the user is looking straight in font, the pixels on both the vertical lines are black. This is interpreted as the "center" direction of the user's eye.

When user looks towards left, the pixels on the left vertical line are black, but the pixels on the right vertical line are white. This can be seen in the Fig. 2 shown below. The closed eye condition is also recognized by the software. This condition is then used to determine the blinking of the eye. The natural blinks of eye are distinguished from the unnatural blinks. The user has to blink his eye for a second if he wants to start moving or stop moving the wheelchair.

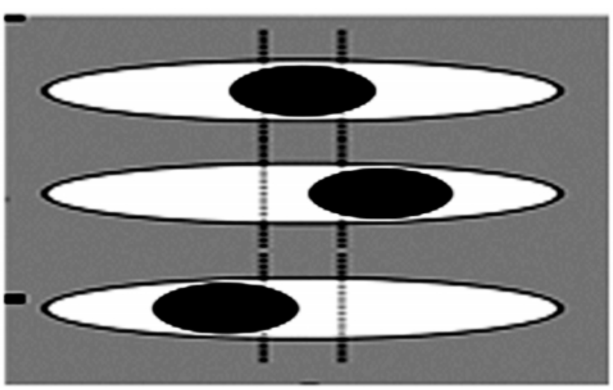

Fig. 2: The Two Scan Lines for Detecting the Motion of Eye

3.1.3 Wheelchair Operation

When the left or right motion of the eye is detected, the wheelchair can be controlled to move in that direction by giving commands to the wheelchair. Theses commands are transferred to the wheelchair using electrical signals which are used the drive the left or right motor of the wheelchair. There are basically two motors connected to the left and right wheels of the wheelchair. The electrical signals are transferred to these motors using some hardware ports, called the communication ports. Generally, the communication port is the parallel port. There are some basic predefined pins of this parallel port which accept the commands given to the wheelchair in the form of electrical signals. For the purpose of demonstration of wheelchair movement using eye motion, a wheelchair model is designed in this project, which works on batteries. This model wheelchair is shown in Fig. 3 below.

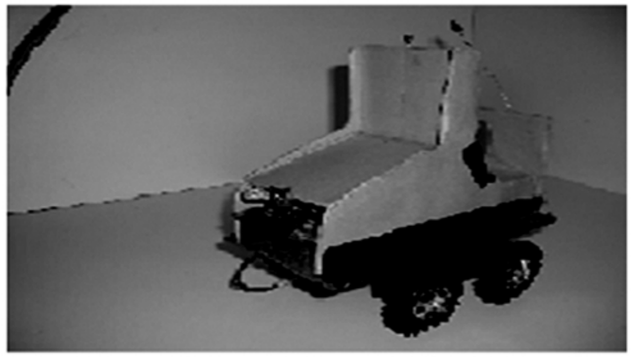

Fig. 3: Model Wheelchair

Four wheels are used in the wheelchair for proper balancing. The movement of wheels is controlled by DC motors which are attached to the wheelchair. Two wheels located on left side of the wheelchair are controlled by one motor and similarly the wheels on the right side are controlled by the second motor. The motors used in this system are small powered, geared DC motors which are generally used in toy cars. The other circuitry built into the wheelchair includes the transmitter and receiver circuits and the obstacle detection circuit. The obstacle detection circuit can be seen in Fig. 2 in the front portion of the wheelchair and the DC motors and the receiver circuit is mounted under the wheelchair. The signal transmitter from the system has an antenna for better range. The commands are transmitted from the image processing software to the parallel port. The 25-pin parallel port connector receives the command in the form a binary number. Based on the binary number received on the pins, voltage is generated on the pins, due to which the transistors connected to the parallel port connector are switched on or off. There are four transistors, one for each direction i.e. left, right, forward and backward. Then, the corresponding signal is transmitted from the transmitter in the circuit to the receiver circuit. Also, an obstacle detection module is added to the system for safety of the user. It involves two IR signal emitters which emit IR signals continuously.

When some obstacle appears in front of the wheelchair, these IR signals are obstructed, and reflected back. These reflected signals are then detected by the IR sensor present just at the side of the emitters. As the IR signals are detected, a circuit is connected to the buzzer, and the buzzer beeps. At the same time, signal is transmitted back to the image processing software in the system so as to stop the wheelchair.

3.1.4 EXPERIMENTAL RESULTS

The system is tested on equal terrain and indoor environment. The USB INTEX night vision camera is attached to the cap which is worn by the user. The camera starts capturing as soon as it is plugged in. Then, we have to start executing the VB program so that the user interface is displayed on the screen. It can be seen in Fig. 6, at the top left corner, there is a list box which displays the list of camera driver objects. There is only one camera driver listed in the list box. First, we select the camera driver from the list and then click "Start Camera View". As a result, the video captured by camera starts displaying in the rightmost picture box. Then, "Start eye detection" is clicked. As a result, the same video starts displaying in the first picture box. At this time, the image processing program has been started and the video in second picture box is the continuous input to the program. Now, a small rectangular portion inside the cornea of the eye (first picture box) has to be selected by the user, with the help of the mouse. It can be seen as a red portion in Fig. 6. The vertical edges of this small rectangular portion form the two scan lines of the coherence algorithm. From the frames extracted, the coherence algorithm determines the direction of eye gaze. In Fig. 6, the position of the eye is seen as left in the label below, because the pixels lying on left scan line are black.

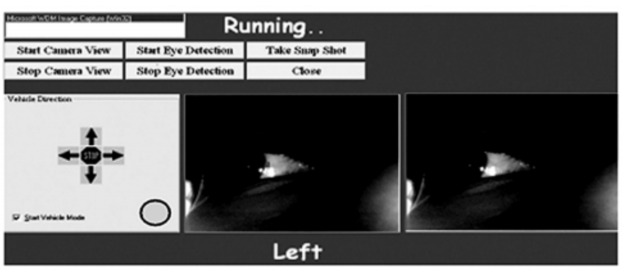

Fig. 4: Eye Motion Detected as Left and Wheelchair is in Running Condition

The starting and stopping of the wheelchair is also controlled by blinking of the eyes. When the user blinks his eyes for a second, then the wheelchair starts or stops moving. That is, if the wheelchair is in moving condition, then after a blink, it stops and vice versa. The direction arrows shown in the figures below are used for testing purpose at the time of demonstration. Beside these arrows, a circle is seen with green color. When some obstacle is detected in front of the wheelchair, this circle becomes red, a buzzer beeps and the wheelchair stops.

Note that this circle doesn't appear red for natural stopping of wheelchair by blinking. Table 1 below shows the test results for wheelchair movement. The Wheelchair and its circuitry required two 1.5 V batteries and three 9 V batteries. The camera used, is a USB camera which can be directly plugged in and used. The camera drivers should be loaded in the system before use.

Table 1
Test Results for Wheelchair Movement

Item	*Capacity*
Frame Capturing Speed (of camera)	20 fps
Maximum Speed of wheelchair	5 km/h
Braking distance	1 to 2 feet (at 5 km/h)
Minimum Radius of rotation	1 feet

3. 2 Interface through Hand Gestures

Instead of managing a huge database, we are going to divide the frame into different scan lines and just examine the pixels which fall under a particular scan line. This reduces the size of database to some extent because we are going to store the coordinates of the scan lines in the form of database. Initially we are going to examine whether this plan work for the color of arm (i.e. the color of shirt) and if it is successful then we are going to implement this for entire hand posture.

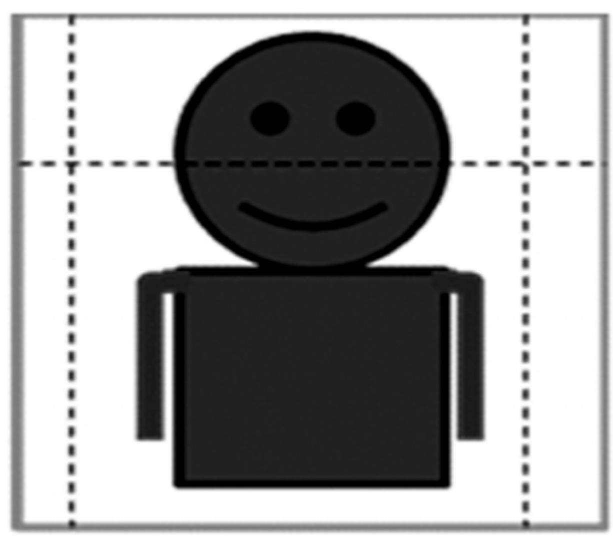

Fig. 5: Initial Position

As shown in Fig. 3, the system scan the pixel color which fall under lines indicated as *L* for left hand movements and *R* for right hand movements respectively and accordingly takes the respective action.

3.2.1 Processing of Scan Lines

Following steps are used for scan lines

- **Capturing Camera View:**

First of all we find the list of driver in the project to start the capturing of the camera view in the picture box. We need to pass the driver detail to this function that is name of the driver and the ID of driver. This function will return a window object on successfully capturing the camera window.

- **Getting Current Frame out of it:**

Now we are having the camera view but the problem is that this view is handled by the O.S. and the camera driver can't work on it for processing. So we need to convert the live video in to processing format. So to get the current view as image we used a function which will convert the view in to image format which be processed by our project. Next we used a variable which can hold entire image in memory. After this step we have image in variable which we can used to extract the pixel.

- **Finding Pixel RGB value:**

As we are having image in variable we have pixels not pixel RGB information. So we declare one variable which is of type Byte and by using some function we can get the pixel RGB value detail which in range of 0 to 255. Here we provide an Image Data ( ) 3D array to hold pixel detail. After execution of this function we get the array filled with pixel detail where to get the *R* value of any pixel we can use *R* = Image Data (2, *x*, *y*) where *X*, *Y* is the coordinates of the pixel. This is same for all other that is *G* and *B*

- **Comparing Pixel color:**

After extracting the pixel value we compare the value and find out which color it represent.

Eg. If value for (*R*, *G*, *B*) is (255, 0, 0) then it is Red color. Similarly (0, 255, 0) is Green and (0, 0, 255) is Blue.

- **Decision making:**

Depending upon the color value of pixel we extract the value of pixel under specific scan lines by using get near by value () function. This function take the value for red color in initR variable and image data () array as argument. The function returns a Boolean value which shows the pixel in current red, green and blue region and if it is found then set red color line in that region. The function picks a pixel and compares other pixel with current pixel, if value of new pixel is near to current pixel then it ignore that pixel otherwise save this value in array which is used to draw scan lines.

The overall procedure is done in screenshot below. As shown in Fig. 3 the image, here the live image capture by camera is divided into scan lines and then area near each scan line is examined for specific commands. For example, when left hand is moved, area under left scan line is cut and hence it is detected that left hand is moved and same for other movements also. This is treated as commands for robot or any device which then work accordingly depending on the commands given to them.

Fig. 6: Screenshot of Processing Image

4. CONCLUSION AND FUTURE SCOPE

This Paper describes two algorithms, one for interfacing through eye and other for interfacing through hand gestures. Depending on the kind of disability the disabled person can choose the type of interface. Fusion of both the techniques can be done in future.

This paper reports a size and position invariant human posture recognition algorithm. The image is first acquired using an address event temporal difference image sensor and followed by a bio-inspired hierarchical line segment extraction unit. A simplified scan line algorithm is used to get the command from user. The proposed algorithm achieves 88% average recognition rate while features 10 – 100 × computational saving as compared to conventional approach.

So in our propose system we are going to use same algorithm which may be further improved by using dynamic images using CCTV camera, instead of storing all posture images into database and then perform the same action on real-time images as explained in this algorithm.

The proposed wheelchair system is easy to operate by the user. The cap worn by the user is light weighted carrying only a small camera with LEDs. The user has to only look left or right to move the wheelchair towards the desired direction. The diagonal motion is achieved when user looks left or right for only small duration of time. The experimental results are satisfactory which are shown in table 1. The actual implementation of the project requires changes in the wheelchair construction.

REFERENCES

[1] Naouki Kubota, Yu Tomioka, Toru Yamaguchi, (2008). "Gesture Recognition for Partner Robot based on Computational Intelligence", IEEE 2008.

[2] S. Stoddard, L. Jans, J.M. Ripple, and L. Kraus, "Chartbook on Work and Disability in the United States, 1998", *an Info Use Report, U.S. National Institute on Disability and Rehabilitation Research*, 1998.

[3] L. Fehr, W. Edwin Langbein, and S.B. Skaar, "Adequacy of Power Wheelchair Control Interfaces for Persons with Severe Disabilities: A Clinical Survey", *J. Rehabil. Res. Develop*, **37 (3)**, pp. 353-360, 2000.

[4] V. Kumar, T. Rahman, and V. Krovi, "Assistive Devices for People with Motor Disabilities", *in Encyclopaedia of Electrical and Electronics Engineering*, J.G. Webster, Ed. New York: Wiley, 1999.

[5] R.C. Simpson, "Smart Wheelchairs: A Literature Review", *J. Rehabil. Res. Develop.*, **42**, pp. 423-436, 2005.

[6] Joyce Y. Chai Shaolin Qu, "A Salience Driven Approach to Robust Input Interpretation in Multimodal Conversational Systems", Computer Science and Engineering, Michigan State University, East Lansing, MI 48824.

[7] Shoushun Chen, Berin Martini, and Eugenio Culurciello, "A Bio-inspired Event-based Size and Position Invariant Human Posture Recognition Algorithm", IEEE 2009.

[8] LIU Zhe LI Xiao-jiu, "Image Abnormal Region Recognition with Fuzzy Clustering Based on Multi-characteristic Variable Window", *First International Workshop on Education Technology and Computer Science,* 2009.

[9] A.H. Muhamad, Amin, R.A., Raja Mahmood, and A.I. Khan, "Analysis of Pattern Recognition Algorithms using Associative Memory Approach: A Comparative Study between the Hopfield Network and Distributed Hierarchical Graph Neuron (DHGN)", IEEE 8th International Conference on Computer and Information Technology Workshops.

[10] Tong Wen Wang, Lin Gaun, Yao Zhang, "A Modified Pattern Recognition and its Application in Power System Transient Stability Assessment", IEEE 2008.

[11] B.A. Kitchenham, "Procedures for Performing Systematic Reviews", *Technical Report TR/SE-0401, Keele University, and Technical Report 0400011T.1, NICTA, 2004*.

[12] Gyu Myoung Lee, Jun Kyun Choi, Noel Crespi, "Object Identification for Ubiquitous Networking", ICACT 2009.

[13] Tutorials on "Human Robot Interaction" from Wikipedia Free Encyclopedia.

[14] J. Eisenstein, and C.M. Christoudias. "A Salience-based Approach to Gesture-speech Alignment". In Proceedings of HLT/NAACL 04, 2004.