Double Modality Computer Interface for Learners with Special Needs

Habib Hamam[1], Daniel LeBlanc[1], Yacine Ben Ahmed[2], Sid-Ahmed Selouani[2], Yassine Bouslimani[1]
University of Moncton, [1]Moncton Campus, [2]Shippagan Campus, NB, Canada E1A 3E9

***Abstract*—We propose a cost effective human-machine interface interpreting the user's hand or head gestures. This system targets people suffering from reduced mobility and can be used in Computer and Web-based learning. This design allows using simple head movements to perform basic computer mouse operations, such as moving the mouse cursor on a computer screen. To improve functionality, the system uses voice recognition to execute repetitive movements, such as clicks and double-clicks. We will explain two different infrared sensor layouts and discuss the way our design detects and processes a user's head or hand movements. Furthermore, we will explain the benefits and drawbacks of using voice recognition in our design. The conclusion presents our results as well as future perspectives for our design.**

***Index Terms*—Aids for the handicapped, Computer interfaces, User interfaces, Infrared detectors, voice recognition.**

I. INTRODUCTION

A new technique is offered to people with particular constraints to have access to television, personal computers (PC), the Internet and various other instruments of daily personal or professional use. Several applications are targeted. However, in the present work, special attention is given to the minority, suffering from reduced mobility, and to the requirements of professionals such as surgeons. In the operating room, these doctors need to manipulate medical instruments, to consult databases, to display and handle curves or graphs, to activate a program on the PC, etc., while their hands remain inaccessible (either occupied or because of sterilization).

We aim at a cost effective solution for a human-machine interface. By using sensors to detect simple hand or head gestures as well as voice recognition, our solution should offer convivial access to machine control interfaces and communication networks for populations with special needs (people with reduced mobility), for practicing professionals (surgeons, anesthetists, etc.) and for employees working in hazardous situations.

Particular attention is given to the design of a user interface enabling the elegant control of a PC mouse (move and click) by using subtle head or hand movements as well as voice commands. The proposed design is based on infrared distance measuring sensors which detect the head movements and convey that information to the PC by means of a microcontroller and a USB connection.

This article includes a quick review of different gesture based human-machine interaction (HMI) methods already in use. Common disadvantages of existing methods will be exhibited. We then propose our method by first comparing two different sensor layouts and their inherent advantages and disadvantages. We then explain the current hardware and software implementations of our method. The results obtained from the prototype device are then discussed. Improvements to the current design as well as perspectives are presented in conclusion.

II. EXISTING METHODS

Existing methods for gesture based HMI or human-computer interaction (HCI) include vision based tracking, eye tracking, head pointing, switch-related systems and brain computer interfaces (BCI) [1]. Each method shows advantages and disadvantages. Common drawbacks include:

1. The need for specialized software or drivers to be installed on the host PC
2. Bulky hardware which reduces the flexibility and mobility of the device
3. The high cost of the device itself

In addition, solutions such as vision based tracking systems and eye tracking systems often have variable performance when used under varying environmental conditions (different backgrounds, in direct sunlight, etc.).

Existing methods for gesture based HMI or human-computer interaction (HCI) include vision based tracking, eye tracking, head pointing, switch-related systems and brain computer interfaces (BCI) [1]. Each method shows advantages and disadvantages. Common drawbacks include:

Vision based systems use cameras to detect objects. Images from the cameras are processed so that the movement of specific objects can be traced from one image frame to the next. With vision based systems, assistive technologies for reduced mobility users can be created [2].

Eye tracking systems often refer to systems that track and interpret the gaze from the user's eyes. For example, when applied to HCI, these systems detect where a user is looking on a PC screen and move the cursor to the appropriate position [3]. This approach depends on the user's facial features.

Switch-related systems include devices that can be controlled by a user inhaling or exhaling into a tube. Other switch-related systems include devices such as tongue touch keypads [1].

Brain Computer Interfaces (BCIs) allow the user to control a machine or PC with signals directly from the brain. The use of BCI is interesting, but such devices remain expensive and their use needs to be monitored by qualified people [1].

Head pointing systems include commercially available optical head pointing systems such as HeadMouse and SmartNav [5]. These devices offer users the possibility of a head movement controlled HCI. Such devices are mounted on top of a PC monitor and detect the head movements from a user in front of the monitor. Commercially available devices usually require specialized software to be installed on the host PC.

Voice recognition software lets users control a PC through voice commands. Programs designed and then installed on the user PC translate these voice commands into instructions for the computer. Voice control has advantages and drawbacks like any other existing PC control method.

The possibility of having a large number of possible commands is one advantage of using voice recognition. However, using a large number of voice commands (a large vocabulary) requires a dependable voice recognition system. This would require that the PC be "trained" to recognize the voice of a specific user. This tends to reduce the flexibility of the system. Another drawback of the voice only control systems is their inability to provide the same functionality as a standard mouse (i.e. controlling the mouse cursor directly so as to move it on the screen).

In this article, we present an alternative to existing head pointing systems or voice only control systems. By combining the use of gesture based control (either head or hands) and voice recognition, we can exploit the advantages of both control methods.

Like other optical head pointing systems, our device uses infrared light to track the movements of a user's head. An important part of our device is the layout of the infrared sensors relative to the user's head. Using the same sensor layout, we have the possibility to substitute the infrared sensors for a different type of distance measuring sensor, such as ultrasonic sensors. Our system offers the added possibility of being fixed to a user (such as being attached to the back of a wheelchair) instead of being fixed to a specific PC.

Voice recognition enables executing repetitive operations, such as clicks and double clicks without the need to perform repetitive head or hand movements. By limiting the number of voice commands, we can avoid having to "train" the PC to recognize the user's voice.

III. PROPOSED METHOD

The method we propose is to use low cost infrared distance measuring sensors to capture the movements of a user's head. This head movement sensing system is augmented with a voice recognition program installed on the user's PC. The user's head movements will control a mouse cursor on the PC screen, just like a standard mouse pointing device. The user's voice commands allow performing mouse clicks and other similar operations. The data from the sensors is collected by a microcontroller card and is then transmitted to the host personal computer via a USB connection. The sensors, microcontroller and USB controller used will be detailed in section IV. The voice recognition program installed on the computer interprets the user's voice commands and translates them into commands such as "click" (or "left click"), "right click" or "double click". The voice recognition program used will be detailed in section V. A wireless microphone system transmits the voice commands from the user to the computer sound input.

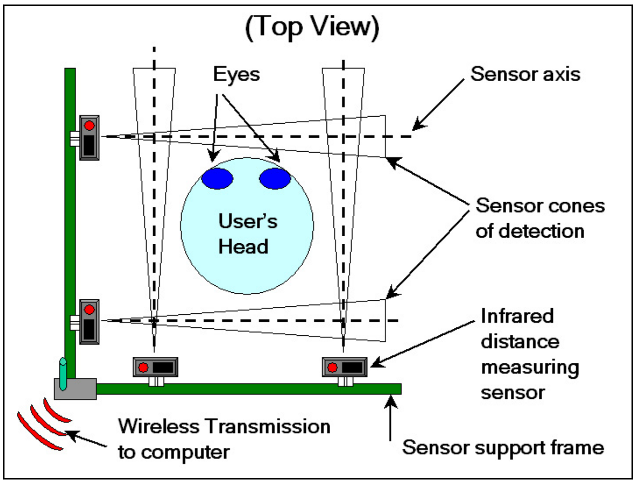

Figure 1. Sensor layout on support frame with four sensors and their positions relative to the user's head.

The layout and the number of sensors used are important parameters for the device's overall performance and cost. We investigated different sensor layouts and opted for the two briefly explained below. For a more detailed explanation of both sensor layouts and their advantages and drawbacks, see [6].

The first sensor layout uses four sensors attached to a holding frame of L-type. The frame has two sides with two sensors on each side as shown in Figure 1.

In this layout, the user's head is located in the neutral position were it is not detected by any of the sensor cones of detection. When the user leans his or her head to either side or towards the front or backwards, the head enters one or two of the sensor cones of detection. The movements are interpreted by the microcontroller and the corresponding mouse cursor actions are sent to the host personal computer.

In the second sensor layout, the number of sensors is reduced from four to two. The locations of the two sensors are shown in Figure 2.

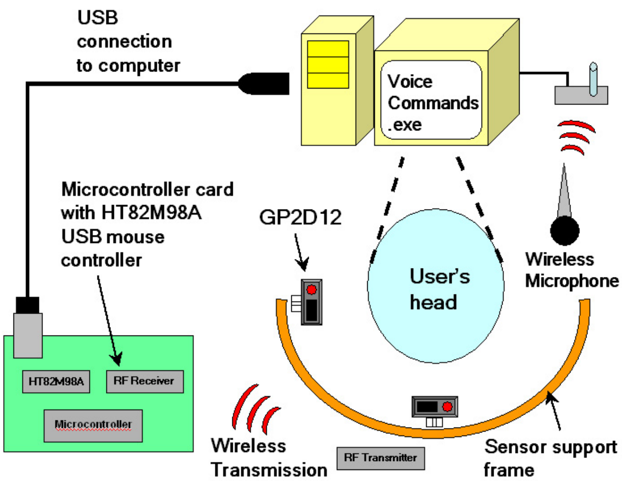

Figure 2. The two sensors layout.

In this sensor layout, the user's head is located directly in the sensor cones of detection. This layout requires that the infrared sensors used be able to measure the distance between the user's head and the sensor. When the user leans his or her head backwards or forwards or from side to side, the appropriate sensor will measure the distance between the user's head and the sensor. The microcontroller will interpret the difference in consecutive measurements from the sensor and will convert the head movement to the corresponding mouse cursor movement.

One particular advantage of the layout using two sensors is the possibility of changing the cursor speed using hardware solutions. For example, the system can be set so that when the user leans his head slightly, the cursor moves slowly in the appropriate direction. Gradually, as the user leans his head more and more, the cursor can be accelerated. For users with good head mobility, variable cursor speeds allow better performance (more productivity, etc.).

IV. HARDWARE IMPLEMENTATION

The method we propose uses infrared distance measuring sensors to detect the presence of a user's head. The sensors are connected to a microcontroller based printed circuit board which is responsible for acquiring and processing the data provided by the sensors. The USB interface is controlled by a dedicated chip (Holtek HT82M98A USB+PS-2 mouse controller) that communicates directly with the personal computer to which the device is connected. The device uses standard USB drivers which are installed by default in most operating systems such as Windows XP.

To allow greater flexibility and mobility, we have created a simple wireless RF link between the sensor support frame and the computer. A microcontroller card on the sensor frames transmits the necessary data to the USB Mouse Controller through a wireless link (Fig.1 and 2).

A. Infrared distance measuring sensors

The infrared sensors used for this design are the GP2D12 General Purpose Distance Measuring Sensor manufactured by SHARP. The cost of each sensor is around 8 $US. These sensors have a range between 10 and 80cm. This sensor detection range is adequate for our intended purpose. Each sensor emits an infrared beam which is then reflected by objects in the sensor cone of detection. This cone is narrow enough to limit interference from objects around the sensor. The GP2D12 provides a voltage that is proportional to the measured distance from the reflecting object. In our case, we use the sensors to measure the distance from the sensor to the user's head or hands. For a more detailed explanation of the GP2D12 and its capabilities, see our previous work [6]

B. Sensor data capture

The infrared distance measuring sensors are connected to a microcontroller based printed circuit board. This connection can either be wired or be in the form of a wireless RF link as explained above. The microcontroller on this card is responsible for acquiring data in the form of analog output voltages from the sensors. The microcontroller must also process the acquired data to evaluate its validity. After a movement has been confirmed, the microcontroller must communicate with the USB mouse controller located on the same printed circuit board.

The microcontroller used in the design is the ATmega8 manufactured by Atmel. This microcontroller possesses a large number of configurable input and output ports and an analog-to-digital converter (ADC). The performance of the chosen microcontroller is more than adequate for the proposed usage.

The analog output voltages from the sensors are converted to an 8-bit digital value using the microcontroller's ADC. So as to avoid false detection of objects due to errors during the analog-to-digital conversion or due to errors at the sensor level, 5 consecutive values are converted and stored in the microcontroller. If 3 or more of these values indicate that a movement has occurred, then the movement is confirmed and appropriate actions are taken by the microcontroller.

Once a movement or a click has been confirmed by the microcontroller, this information must be passed on the USB mouse controller which will then convey the relevant data to the host personal computer. Signals from the microcontroller are formatted so as to be compatible with what the USB mouse controller expects to receive.

C. HT82M98A USB Mouse Controller

The USB mouse controller used for this device is the HT82M98A USB+PS/2 Mouse Controller from Holtek. This controller is not only able to make a USB connection with a PC, but also a PS/2 connection. Designers have the choice of using either a USB connection or a traditional PS/2 connection. This adds to the compatibility of the device since older personal computers may not have the required USB ports or USB default drivers.

The HT82M98A is a typical USB mouse controller used in commercial designs for ordinary mice. It is the HT82M98A USB controller's inherent compatibility with personal computers that allows the device to establish a standard connection to the host computer without requiring extra hardware.

The left, right, up and down movements are controlled using the X1, X2, Y1 and Y2 inputs on the HT82M98A. The X1 and X2 (as well as the Y1 and Y2) inputs are two lines connected to the microcontroller on which square waves are transmitted. The square waves on line X1 are shifted in relation to the square waves on line X2. This is similar to how standard ball and wheel mice work. In a standard ball and wheel mouse, the offset square waves are created by a combination of wheels with holes, LEDs and phototransistors for the X and Y axes. These offset square waves are the means by which the HT82M98A counts the movements on the X or Y axes. The HT82M98A also has inputs for the left, middle and right buttons of a mouse. In our design, the click operations are controlled through voice commands and a program running on the host computer. The HT82M98A thus does not control any clicks in our design.

D. Connection to Personal Computer

Using the HT82M98A mouse controller, we are able to have either a USB or a PS/2 connection to the PC. We have chosen to implement a USB connection with the computer since the computer's USB ports can provide sufficient current (up to 500mA) to power the HT82M98A, a microcontroller and the wireless receiver. The microcontroller located on the sensor frame, as well as the sensors, are powered by a set of batteries. The device uses the Human Interface Device (HID) driver which is part of the default tools in Windows XP to control pointing devices such as USB mice. The HID driver can be installed on other versions of Windows who don't have the driver installed by default.

E. Wireless RF Link

In our current design, we are using basic RF transmitters and receivers with no predefined protocols. The components used are manufactured by Linx Tech. Inc. These transmitters and receivers integrate everything needed in one convenient package and are available in different frequencies, such as 315, 418, 433 and the 900MHz band. The only outside component needed are antennas. We are currently testing alternative RF components so as to lower the overall cost of the device.

V. SOFTWARE IMPLEMENTATION

The software program developed is a small program (around 720 KB) which interprets a specific set of voice commands from the user. These commands are used to execute repetitive actions such asclicks. For the program to work, the Win32 Speech API (SAPI) needs to be installed on the computer. We have Microsoft Speech SDK 5.1 [7] installed on our machines which provides everything necessary to run our voice command program. The commands available to the user are the following: "left click" (or just "click"), "double left click" (or just "double click"), "right click", "double right click", "middle click". The user can also keep a button pressed so as to highlight a group of objects. The command "left down" keeps the left mouse button pressed until the "left up" command is given. This set of available commands allows executing meaningful tasks on the computer since all the main mouse click operations are available.

Our software program uses generic speech recognition. This means that the computer doesn't have to be "trained" to understand a specific user's voice. This allows greater flexibility, since the system is accessible to new users without the need to "train" the computer.

Figure 2 shows the device layout with the hardware components described in the previous sections and the appropriate connections. The connection between the sensors and the microcontroller is wireless. The microphone is also wirelessly connected to the voice interpreter. Hence the whole user terminal (bottom right part of Figure 2) is wirelessly connected to the fixed systems. Thus, the user terminal can be moved freely.

VI. EXPERIMENTS

With the layout using two or four sensors and the hardware and software described above, we have tested the performance of the device. The following actions can be conducted reliably using the device:

1. Left, right, up and down movements of the mouse on the screen. With the layout using two sensors, the speed with which the cursor moves on the screen depends on how far the user tilts his head to either side or forwards and backwards. The device was tested with three different cursor speed levels. The device interpreted the user's intentions correctly at the different speeds.
2. Mouse clicks and other click related operations. The user's commands are correctly interpreted by the computer which is not necessarily "trained" to the specific user's voice. Wrong words and altered pronunciations are merely ignored and launch no operation. Repeating the command with the correct vocabulary and pronunciation corrects the situation.

Using the infrared sensors described (GP2D12) also renders our device less susceptible to variations in performance due to a user's skin colour or hairstyle. The infrared beam of the sensor is best reflected on bright white surfaces, but the attenuation of the reflection caused by darker colours is not significant enough to affect the performance of the sensors in our application.

VII. CONCLUSIONS

In this article, we have discussed the design of an infrared-based HMI device. We have used this design as a human-computer interface able to replace a traditional computer mouse. We have advanced two different infrared sensors layouts using a combination of two or four sensors each. Given that the four sensors layout includes a neutral zone, it is more comfortable for the users and tolerates more inaccurate movements. A description of the hardware used in the implementation of the design has provided incite into how our device works. The software aspects of the device have also been explained. Finally, we have done preliminary tests on the design so as to insure that the working concept is sound.

Existing gesture based solutions are briefly reviewed and advantages as well as disadvantages associated with these solutions are pointed out. Our approach has four main assets. First, it allows HMI without need of accurate movements (four sensors layout), which is especially valuable for the population with special needs. Second, repetitive actions are performed by voice commands and not by gestures which are generally awkward when they are repeated frequently. The system is robust since in terms of gestures it allows inaccurate movements and in terms of voice recognition it uses a limited vocabulary. It is thus easily distinguishes between the few voice commands. Fourth, by providing a neutral zone, this approach implements an easy way of defaulting to stand-by mode (no operation).

Further work on the project will include improving the reliability of the system as well as integrating new technological advances. The wireless RF link is still prone to interference. The main problem with the design of a reliable RF link is the manufacturing process used to create printed circuit boards in our laboratory. The production method of our circuit boards combined with the relatively small size of most RF components results in poor performance. Advances in wireless sensor networks architectures and protocols create new areas of application for wireless sensor networks. A wireless sensor network created between each sensor and the host computer would permit a more flexible & portable system. We aim to apply wireless sensor network progresses so as to benefit the design of HCI.

Constant use of certain muscle groups, such as the neck muscle supporting the head, can lead to fatigue [5]. An ergonomic study of the permissible head movements within the sensor detection area would reduce the possibility of muscle fatigue which could eventually lead to chronic health problems. This ergonomic study would be necessary to ensure that a final design doesn't expose the user to undue health risks.

REFERENCES

[1] T. Surdilovic, Y. Zhang, Convenient intelligent cursor control web systems for Internet users with severe motor-impairments, International Journal of Medical Informatics, vol. 75, January 2006, pp. 86-100.

[2] J.D. Crisman, M. E. Cleary, J.C Rojas, The deictically controlled wheelchair, Image and Vision Computing, vol. 16, 1998, pp. 235-249.

[3] Computer Vision and Image Understanding, vol. 98, Special issue: eye detection and tracking, 2005, pp. 1-210.

[4] C.H. Morimoto, M.. Mimica, Computer Vision and Image Understanding, vol. 98, Special issue: eye detection and tracking, 2005, pp. 4-24.

[5] R. Ortega, Unusual access methods, in Proc. CSUN's 19th An. Int. Conf. on Tech. & Persons with Disabilities, Los Angeles 2004: http://www.csun.edu/cod/conf/2004/proceedings/143.htm

[6] D. LeBlanc, H. Hamam, Y. Bouslimani, Infrared-based Human-machine interaction, Proc. of the ICTTA, 2006.

[7] MS Speech SDK5.1: www.microsoft.com/speech/download/sdk51

AUTHORS

H. Hamam is with the University of Moncton (Canada) and Canadian University of Dubai (UAE), (e-mail: Habib.Hamam@umoncton.ca).

D. LeBlanc is with the University of Moncton, campus Moncton (Canada) (e-mail: leblancd@umoncton.ca).

Y. Ben Ahmed is with the University of Moncton, campus Shippagan, (Canada) (e-mail: yacine@umcs.ca).

S.A. Selouani is with the University of Moncton, campus Shippagan (Canada) (e-mail: selouani@umcs.ca).

D. LeBlanc is with the University of Moncton, campus Moncton (Canada) (e-mail: Yassine.Bouslimani@umoncton.ca).

Manuscript received 22 May 2007. This work was supported in part by the NSERC - Canada. Published as submitted by the author(s).

Published as submitted by the author(s).