International Journal of Enhanced Research in Science Technology & Engineering, ISSN: 2319-7463
Vol. 3 Issue 5, May-2014, pp: (163-169), Impact Factor: 1.252, Available online at: www.erpublications.com

Efficient Gestures and Voice Based Computing using Hardware Implementation

Azam Rafique Memon[1], Tarique Rafique Memon[2], Khuhed Memon[1]
[1]Department of Electronic Engineering, Mehran UET, Jamshoro, Pakistan
[2]Department of Electronic Engineering, QUEST, Nawabshah, Pakistan

Abstract: **This research is focused on solving speed and accuracy problems that current Human-Computer Interface (HCI) fights with. The system uses a separate hardware and multimodal interface to improve system performance and reduce error rate. By using multimodal interface, the gestures and speech both will be used to interact with computers. With the help of hardware circuitry we can improve the speed of Gestures based computing, hence reducing gestures recognition and processing load from computer's microprocessor. The DSP Processor will handle all the processing of recognition of gestures and will issue commands to computer via PS/2 or USB Ports. Focus of this research is given to control mouse pointer and the features related to it in most of the applications.**

Keywords:computer control; DSP processing; gestures detection; Human-Computer Interface (HCI); postures detection; speech recognition.

Introduction

Human-Machine interface (HMI) is considered as the key factor in development of controlling the computer efficiently. Many interfaces have been developed over years to control different applications [1]. Recently, the research is mainly focused on oral computing, and specific applications have been controlled through gestures and speech, like Earth 3D [2], Gaming [3] and Robots [4]. But the work on whole operating system is not done yet. Along with this the speed and accuracy is still an issue in gesture based computing.

This research is focused to improve performance of interaction with computer by using hardware based acceleration instead of software algorithm running inside the computer. And to reduce error rate by using multimodal interface [2]. The use of gestures and speech can help improve accuracy. Many algorithms have been developed to perform all mouse actions, which will help to access whole computer instead of just specific application.

There are many DSP Processors, which can be programmed through MATLAB codes. These processors will help us to perform all related work of gesture recognition and issue specific commands to computer. This will also help to control mouse by connecting the output from processor to PS/2 port on computer. The use of this technique will reduce overload from computers microprocessor and improve efficiency.

In noisy environment, gestures will help to interact with computer, while people who are disabled or encumbered can use speech [2].

Related Work

Mayron W. Krueger firstly proposed a new form of interaction between human and computer in 1970s [5], [6]. Since then, scientists and engineers have started their research work in this field. Most of them have worked on controlling specific software [2], [7]. Very few have worked on designing a general system to control over all work. All have worked on computer programming [8]–[13] to detect gestures or postures, for that they have used Matlab [8]–[10] and LabView [13] and few have used Sensors [7], [14], [15], but again all processing is done by the computer's microprocessor, which increases overhead to computer's microprocessor and resulting in slowing down overall system performance.

Our previous work [9] was on gesture recognition using color bands [9], [12], which was simulated on Matlab. We have used a simple color detection technique, which worked efficiently. The problem with this and the algorithms used by others [1], [4], [5], [8], [11], [14], [16], is speed and accuracy. In our previous work, the system had good accuracy under one condition, i.e if background has no red or green colored object, but with this the overall computer performance slows down.

Several improved algorithms have been used by others like edge detection [5], Histogram [4], skin color detection [1], [5], wavelet transform [8], [11], algorithms using Neural Network [14] and Fuzzy Logic [16] etc. The only problem with these was speed and accuracy.

Most of the work was done on computer based software, which run on computer itself. Krum D.M. [2] has controlled Earth 3D using multimodal interface (i.e. Gestures and speech). Dr. Jane J. Stephan [5] used different gestures, postures and used edge detection and segmentation technique to detect skin color from complex backgrounds. Triesch, J [8] proposed ways to

detect hand postures against complex backgrounds using wavelet transform. Kjeldsen, R. [17] worked on head gestures instead of hand gestures or postures, and detected facial expressions and head movements. Frigola, M.[1] used several cameras to detect head and hand gestures and used color detection, motion detection and segmentation techniques to recognize gestures. Ionescu, Dan. [3] used many utilities to detect gestures including pulse-laser-light, gain modulation camera and phase-lock loop to detect depth of object and gestures and reconstructed them virtually. Koceski, S. [4] took histogram of acquired image, by taking horizontal and vertical histogram the position of hand can easily be detected, this was an efficient algorithm for posture detection. Lech, M. [16] used Fuzzy logic to recognize gestures and used distance based motion detection technique and worked on white boards by using projector, the problem in this research was accuracy, and slow performance of the designed system, as well as overall system performance is also affected.

The hardware was also used by few, in which they have used some sensors like Kinect [7], [15] and Flex Sensors [14]. But again this input is fed to computer for further processing. The computer keyboards are already been designed that have some extra functions like opening calculator, my computer, browser etc. These keys have specific codes defined, when code is sent via PS/2 and USB, the pre-defined action is performed.

Proposed System

We have extended our previous work, to increase computer performance and accuracy. Our system is based on color detection, it detects two color bands Red and Green, which we had worn on finger and thumbs of both hands and then according to the color positions and gestures, some pre-defined actions are performed. The mouse pointer follows red color and green was used for support in performing some actions and operations. When distance between red and green color band decreases, the mouse click action is performed and when distance increases, mouse release operation is performed.

The proposed system is completely based on hardware circuitry. The system uses a single camera, microphone, DSP processing kit, speech recognition circuit, PS/2 interface and USB interface. The system uses two techniques to operate computer, first technique is using Gestures/Postures and the other one is using Voice commands, both are implemented using hardware. This multimodal [2] interface helps to interact with computer more efficiently. The use of hardware increases accuracy by decreasing load from computer's microprocessor. The gestures and posture recognition is done by DSP processing kit, which directly produce pulse similar to that a PC mouse controller uses to perform mouse functions. Another feature is the use of speech to perform computer operations, by using HM2007 [18], which recognize speech and issues commands to perform particular tasks.

The DSP processor is directly be connected to computers PS/2 port, to perform mouse functions. And the other commands to be issued by DSP kit and by speech recognizer will be connected to computer via USB port.

The speech recognizer will act like a keyboard with advanced function keys to open browser, My Computer and other controlling functions. The software that will be controlling this hardware will only be its drivers, which will perform actions according to instructions issued by this system. This will consume less processing power of computer than that of image acquisition, detection, segmentation and recognition of gestures and recognition of speech. All such load will now be handled by our designed hardware system, which has its own processor.

Figure 1. Shows the block diagram of proposed system:

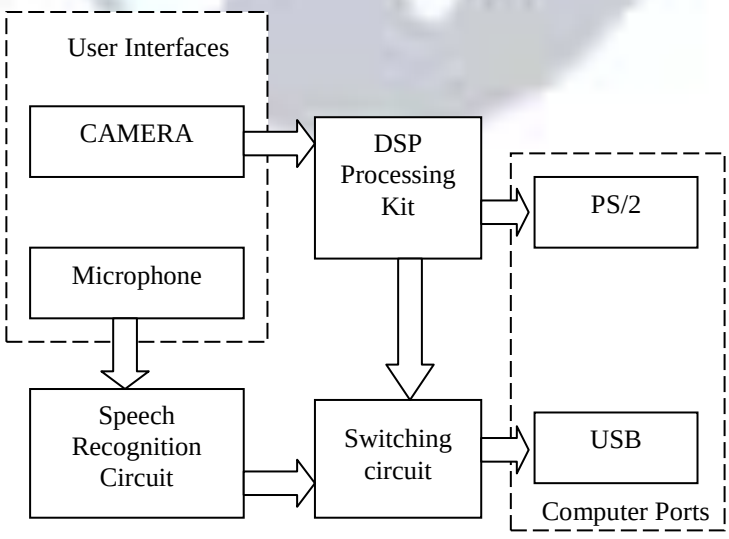

Figure 1: Block Diagram of System

Gestures and Postures

Many gestures and postures recognizing algorithms have been designed in this system. The gestures are used to perform mouse functions, while postures perform different operations in applications, which are common in most of windows based applications, like Open, Close, Exit etc

A. Postures:

The postures are also called as static gestures. The posture recognition is done by using simple self-designed algorithm based on current position of color-bands and distance between color bands.

Our target is not the method used for this, but our purpose is to improve performance and accuracy by using hardware instead of software. So we have used simple techniques, just to make sure that it is possible using hardware and it improves system performance as well.

B. Gestures:

It detects the moving objects, and considers the objects that are skin colored, and discards all others. Then it detects two colored objects, red and green, then it performs calculations between the detected objects to recognize gestures. Once the gesture is detected, it performs specified action related to that gesture.

The gestures are simply the hands in motion. The system detects all moving objects, and consider only objects that are red and green colored, from all these objects, the mouse pointer follow the red moving object. Similar work was carried out in [9]. But in this work, we are also adding voice commands for better accuracy. The gestures used to perform mouse press and release functions are shown in Figure 2 [9]:

The similar algorithm will be used here, it is much efficient algorithm based on MATLAB programming [19].

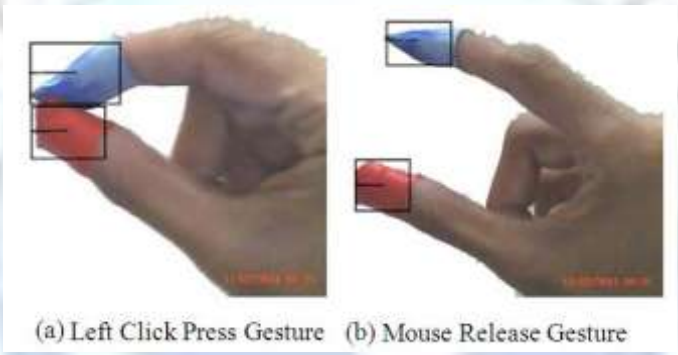

Figure 2: Left Button Press and Release Gestures

Algorithms

Different techniques and algorithms have been used to detect gestures and postures. Reference [20] Proposed efficient algorithms and calculations for motion detection. To detect gestures, the combination of motion detection and edge detection techniques will be highly efficient.

From many algorithms, the simple and efficient algorithms have been taken that can be implemented on DSP board using Matlab.

A. Gesture/Posture Detection:

To detect and recognize different gestures and postures, the acquired image is then passed through many algorithms like: Motion Detection [20], [21], Edge Detection, Segmentation and Histogram [4], [22]. Our previous work on gesture recognition is described in [9], [19]. The general algorithm is described in Figure 3. This algorithm will be repeated infinite times.

B. Mouse Functions:

To perform mouse functions, the recognized gestures are to be converted in appropriate PS/2 mouse signals [23], [24] and fed to computers PS/2 port. Any movement of hand in left direction causes the mouse to move in left direction and so on. When distance between finger and thumb is less than preset threshold values then mouse will make a left click, and mouse button will be released when distance is greater than preset threshold values [9], [19].

C. PS/2 Protocol

When gestures are recognized, the hardware circuit will issue set of commands to PS/2 port[23], to perform articular mouse functions.

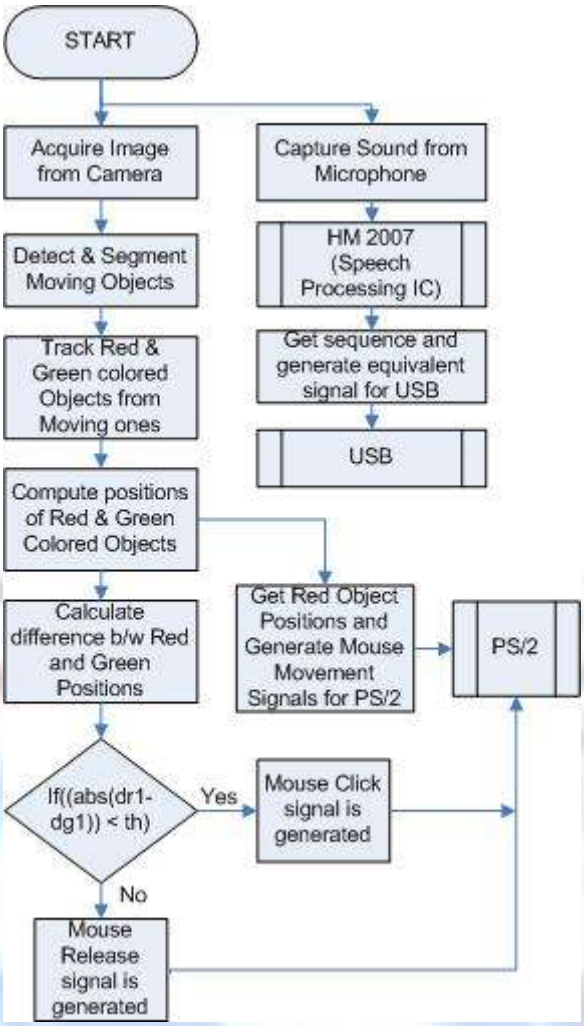

Figure 3: General and Simplified Algorithm of Proposed System

Fig 4 [26], shows block diagram of speech recognizer:

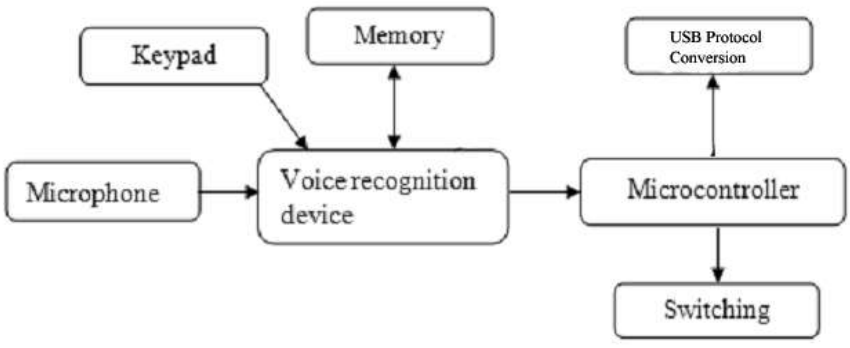

Figure 4: Speech Recognition circuit block diagram

The word recognition flow diagram is shown in Fig 5:

USB monitoring software is designed in visual basic, whose purpose will only be to monitor USB port and perform functions.

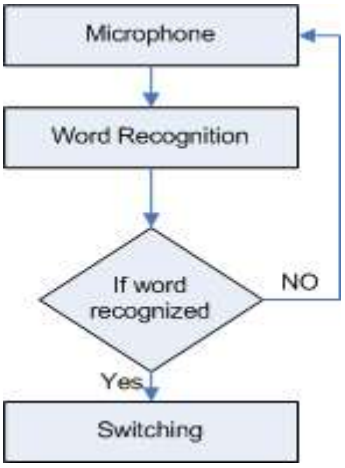

Figure 5: Speech recognition flow diagram

Results

A. **Speech:**

The accuracy of system based on taking 20 trials of each command, as shown in Fig 6:

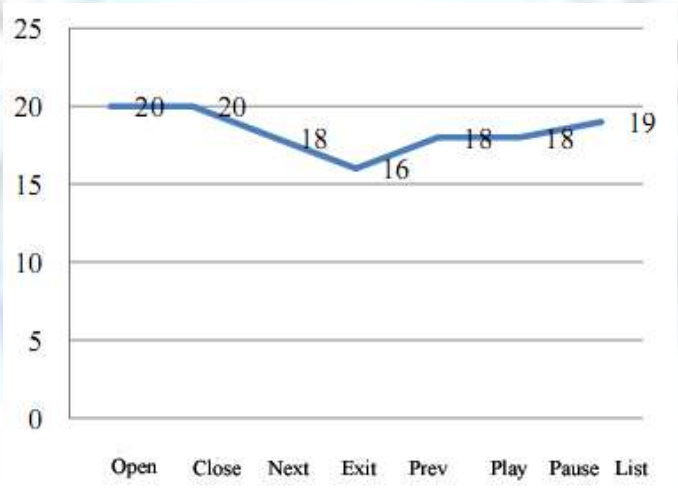

Figure 6: Accuracy Graph of speech recognition circuit

Some mismatching in recognizing commands have been faced, due to similar pronunciation of words, like Next and Exit etc.

B. **Expected Results for Gestures and Postures Recognition:**

The expected accuracy for gestures is 70% to 85%, while the expected accuracy for posture recognition is 75% to 90%.

The expected accuracy for mouse functions is above 90%, with less error rate and high speed performance as compared to other similar systems.

Applications

This area has been a new area of research and it has increasing number of applications:

1. Gesture based Gaming
2. Human Motion sensing
3. Computer Applications
4. Quick operation
5. Mobile Devices

Future Recommendations

In future along with resolving accuracy and performance issues, which may occur in future, this system can be increase to perform other extra functions too. As it can detect different objects in our hands and get information about them from google etc.

For gamming, along with keyboard, mouse and joystick there will be gestures interfacing option. By which users can play games by just performing different actions in front of camera to play game, this will give them a joy as they are playing it in their real life.

For speech recognition, the number of commands will be increased to all windows commands, move, delete, open my computer etc.

Limitations

For gesture and posture recognition, light can be an issue, as camera works on light, so proper lightening is necessary. In colored light gestures and posture recognition system may not work, but in that case speech recognition will surely work.

Speech recognition has a problem in recognition, so user may have to repeat the command more than once.

Acknowledgment

I would like to thank Almighty ALLAH, our parents and Engr. Hafeez Qazi, who gave us this great idea and our supervisors Dr. Bhawani Shankar Chowdhry, Engr. Khuhed Memon and Engr. Attiya Baqai, for their kind help.

Conclusion

There is a new trend to produce and display 3D applications. This type of systems will enhance performance of these applications and will increase user interest to words computing.

Previously the proposed systems were limited to specific applications; none of them have tried to control mouse functions through gestures. As mouse is common in all applications. This system not only control all mouse functions but also allow us to use most common features in most of applications.

The Multimodal interface increases efficiency of the system and use of hardware instead of software program significantly increases system performance.

References

[1]. David M. Krum, Olugbenga Omoteso, William Ribarsky, Thad Starner, Larry F. Hodges, “Speech and Gesture Multimodal Control of a Whole Earth 3D Visualization Environment”, In proceedings of “VISSYM '02 Proceedings of the symposium on Data Visualisation 2002”, 2002.

[2]. Frigola, M., Fernandez, J., Aranda, J., “Visual Human Machine Interface by Gestures”, in proceedings of “International Conference on Robotics & Automation”, 2003.

[3]. Ionescu, Dan; Ionescu, Bodgan; Islam, Shahidul; Gadea, Cristian;, “A New Method for 3D Object Reconstruction in Real Time”, “IEEE International Joint Conferences on Computational Cybernetics and Technical Informatics”, May, 2010.

[4]. Koceski, S.; Koceska, N.;, “Vision-based Gesture Recognition for Human-Computer Interaction and Mobile Robot's Freight Ramp Control”, In Proceedings of 32nd Int. Conf. on Information Technology Interfaces”, June 2010.

[5]. Dr. Jane J. Stephan, Sana'a Khudayer, “Gesture Recognition for Human-Computer Interaction (HCI)”, “International Journal of Advancements in Computing Technology” Volume 2, Number 4, October 2010 .

[6]. “Myron W. Krueger”, http://en.wikipedia.org/wiki/Myron_W._Krueger, [Accessed on June 12, 2011].

[7]. Triesch, J., von der Malsburg, C., “Robust Classification of Hand Postures against complex backgrounds”, In Proceedings of “The Second International Conference on Digital Object Identifier”, 1996.

[8]. Kjeldsen, R., “Head Gestures for Computer Control”, Proceedings. “IEEE ICCV Workshop on Digital Object Identifier”, 2001.

[9]. Lech, M.; Kostek, B.;, “Gesture-based Computer Control System applied to the Interactive Whiteboard”, Proceedings of the 2nd International Conference on Information Technology, ICIT 2010, June 2010.

[10]. Azam Rafique; Abdullah Ansari;, “A New way to use computers with gestures and Images”, In proceedings of 26th IEEEP Student Seminar, Pakistan, 2011.

[11]. Nikhil, S.; Mohan, S.; Ramya, B.; Kadambi, G.R.;, “Design and Development of a DSP Processor Based Reconfigurable Hand Gesture Recognition System for Real Time Applications”, International Conference on Signal and Image Processing (ICSIP), 2010

[12]. HM2007 Datasheet, “http://www.datasheetcatalog.com/datasheets_pdf/H/M/2/0/HM2007.shtml”, [Accessed on June 13, 2011].

[13]. William T. Freeman, David B. Anderson, “Computer Vision for interactive Computer Graphics”, “IEEE Computer Graphics and Applications”, Volume 18, Number 3, June 1998.

[14]. Amat J., Casals A., Frigola F., Pages J.. Possibilities of Man-Machine interaction through the perception of human gestures. In journal Contributions to Science, Institut d’Estudis Catalans, Barcelona , 2000.

[15]. Azarbayejaui A. and Pentland A. Real-time self- calibrating stereo person tracking using 3-D shape estimation from blobs features. In Proceedings IEEE Conference on Computer Vision and Pattern Recognition, Vienna, 1996.

[16]. Brand J. et ali. "A comparative assessment of three approaches to pixel-level human skin-detection". In proceedings of IEEE International Conference on Pattern Recognition 2000, 15(1), pp. 1056-1059.

[17]. Darrell, T.; Gordon, G.; Hawille, M.; Woodfill, J. "Integrated person tracking using stereo, color, and pattern detection". In proc. of IEEE Computer Vision and Pattern Recognition, pp. 601-608, 1998.

[18]. Feraund, R.; Bernier, O.J.; Viallet, J.-E.; Collobert, M. "A fast and accurate face detector based on neural networks". In IEEE Transactions on Pattern Analysis and Machine Intelligence, Volume: 23 Issue: 1, pp. 42 -53, 2001.

[19]. Ohya J., Miyasato T., Nakatsu R. "Virtual Reality Technologies for Multimedia Communications". In Mixed Reality - Merging Real and Virtual Worlds. Ed. Y. Ohta, H. Takamura. Springer-Verlag, NY, 1999.

[20]. Nariman Habili, Cheng Chew Lim and Alireza Moini, "Segmentation of the Face and Hands in Sign Language Video Sequences Using Color and Motion Cues," IEEE transactions on Circuits and Systems for Video Technology, vol. 14, pp. 1086 – 1097, 2004.

[21]. Lae Kyong Lee, Sungshin Kim, Young-Kiu Choi and Man Hyubg Lee, "Recognition of hand gesture to human-computer interaction," IEEE transactions on Industrial Electronics, vol, 3, pp. 2117-2122, 2000.

[22]. Y. P. Lew, A.R. Ramli, S.Y. Kaoy, R. Ali and V. Prakash, "A Hand Segmentation Scheme Using Clustering Technique in Homogeneous Background," Student Conference on Research and Development, pp.305-308, 2002.

[23]. Gestures Controlled Computer, "http://www.facebook.com/video/video.php?v=211686848848463", [Accessed on June 14, 2011].

[24]. "Mouse Connector", "http://www.tavi.co.uk/ps2pages/ohland/mouse.html", [Accessed on June 14, 2011].

[25]. "Mouse (PS/2) connector pinout", "http://pinouts.ru/Inputs/PS2Mouse_pinout.shtml", [Accessed on June 14, 2011].

[26]. Mohd Safuan Bin Abdul Kadir, "Voice Activated Switching Device", May, 2010.