Gesture Controlled Touch Less Response using Image Processing

Ms.Subiya Yaseen,Ms.Shireen Fathima,Mr.Surendra K.V,Mr.Jebran P
Assistant Professor: Department of ECE
HKBK College of Engineering,
Bangalore, India

Ms.Saba Sanober
B.E Graduate: Department of ECE
HKBK College of Engineering,
Bangalore, India

Abstract

In the modern world where almost everything is digitized, different means of interacting with the digital devices are gaining immense popularity and importance. This is because of the advancements of technologies that changed the idea that interacting with the technology is limited to dial pads, keyboards, mouse and touchscreens. Our project is intended to perform various tasks on computer such as capture an image, play an audio, open MS PowerPoint and work with it etc. with the help of gesture recognition and image processing principles only with the help of coloured sensors(caps/LEDs) worn on the fingers of the user. It is a gestural interface that augments the physical world around us with digital information and lets us use our natural hand gestures with coloured sensors to interact with that information. This technology paves a path for enormous applications. This technique enables human interaction with computers in a more direct way without using any external interfacing devices. It can provide a much better alternative to text user interface and graphical user interface [6]to interact with the computer.

Keywords –: Touch less [3], Seventh Sense , Matlab

1. Introduction

Every one of us is aware of the five basic senses – seeing, feeling, smelling, tasting and hearing. These senses have evolved through millions of years. Whenever we encounter a new object/experience our natural senses try to analyze that experience and the information that is obtained is used to modify our interaction with the environment. But in this age of technologies the most important information that helps one to make right decision is something that cannot be perceived and analyzed by our natural senses. That information is the data in the digital form, and it is available to everyone through sources like internet. Our effort is to connect this data in the digital world in to the real world. Although miniaturized versions of computers help us to connect to the digital world but there aren't any device as of now which gives a direct link between the digital world and our physical interaction[7] with the real world. This technology helps to bridge the gap between tangible and non-tangible world.

Our project is focused to implement gestural interface that augments the physical world around us with digital information and enabling us to use our natural hand gestures[4] to interact with that information.This enables to perform various tasks on computer such as capture an image, play an audio, open MS PowerPoint and work with it etc. with the help of gesture recognition and image processing principles with the help of coloured sensors(caps/LEDs) worn on the fingers of the user.

We have a camera which is used as a sensor.The camera recognizes[2] and tracks user's hand gestures[4] using computer-vision based techniques. The software program processes the data captured by the camera and tracks the locations of the coloured caps (visual tracking fiducials) on the user's fingers. The movements[5] and arrangements of these fiducials are interpreted into gestures that act as interaction instructions for the projected application interface. This can be used to implement several applications.

2. Design Methodology

- The camera captures the gestures and sends to image processing block.
- The further processing of the image is carried using MATLAB. For example, different algorithms and conversions are performed to recognize the colours and gestures accurately.
- Different commands for different gestures are pre-defined.
- The appropriate command for the respective gesture is sent for decoding.
- The command is decoded and the output (interpretation of gesture) is obtained.

3. BLOCK DIAGRAM, WORKING AND MODES OF OPERATION

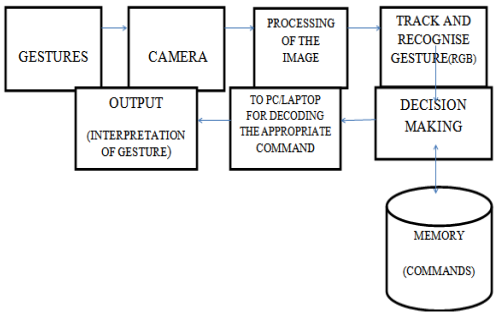


Fig 1:Block Diagram

978-1-5386-1887-5/17/$31.00©2017IEEE

A) Description of the Block Diagram:

The block diagram consists of the following:

- **Gestures:** The hand movements[4][5] with the coloured sensors (RGB) worn on them are the gestures to be captured by the camera. Based on the position of the fiducials (coloured caps) different pre-defined tasks are performed. The position of the fiducials for particular task is also pre-defined.

- **Camera:** Camera is used to capture the gestures. Any camera or the webcam which is able to capture a coloured (RGB) image can be used.

- **Processing of the Image:** Includes different operations that are performed on the captured image to detect the presence of fiducials. This is achieved by performing different operations on the image i.e. image acquisition, conversion from rgb to gray, gray to binary and so on. After the operations are performed relevant algorithms are implemented to increase the accuracy of the detection and for the removal of noise.

- **Track and recognize gesture:** Continuous tracking of the valid gestures is done .If valid gesture is tracked; it is recognized to be valid and is sent for further processing.

- **Decision making:** After tracking of the valid gesture the next step is decision making i.e. what command has to be decoded that corresponds to the respective gesture tracked. Different commands (tasks to be performed on the computer) are stored in the memory priory. Whenever the gesture detected matches one of the commands stored in the memory the corresponding command is fetched and sent for decoding.

- **Decoding:** The decoding of the appropriate command received after decision making is done.

- **Output:** After decoding of the command the task corresponding to the gesture is performed this is nothing but the output. The output is basically the interpretation of the gesture which is destined and defined priorly.

B) Different Modes of Operation

The overall operation and tasks to be performed are characterized into 3 modes based on the similarities the share. The three different modes of operation are as follows:

i) Mode 1: To perform the tasks

First mode of operation consists of basic tasks such as image capture and save, to play an audio, to play a video, open a document, image zoom-in and out etc.

The flow of the operation of mode1 is as follows:

- The camera here is a sensor and is programmed to continuously sense for the gesture and its detection.
- As soon as the image is captured and it is subjected to acquisition for further processing.
- After the image is acquired it is processed by performing different operations on it. The operations include implementation of different image processing algorithms. In our work, algorithms for colour detection (for example, Blob Algorithm) with better accuracy of detection and noise removal are implemented.
- After the implementation of the colour detection algorithm, the colours (RGB) if present are detected and the positions with intensities are obtained.
- If the fiducials (RGB) are present their co-ordinates (positions) are found else the control is transferred where again the camera continuously senses for colours.
- Now the obtained gesture (obtained from the knowledge of position of the colours) is checked for its validity i.e. if the obtained gesture is amongst the pre-defined gestures. If the gesture is a valid one it is recognized else again the control goes back to camera.
- Based on the recognized gesture, the command that corresponds to it is decoded and the task is performed (image capture, audio play etc.).

ii) Mode 2: Working with the PowerPoint

Second mode of operation is for the working with and performing some basic operations on Microsoft PowerPoint presentation software.

The flow of the operation of mode1 is as follows:

- The camera here is a sensor and is programmed to continuously sense for the gesture and its detection.
- As soon as the image is captured and it is subjected to acquisition for further processing.
- After the image is acquired it is processed by performing different operations on it. The operations include implementation of different image processing algorithms. In our work, algorithms for colour detection (for example, Blob Algorithm) with better accuracy of detection and noise removal are implemented.
- After the implementation of the colour detection algorithm, the colours (RGB) if present are detected and the positions with intensities are obtained.

- If the fiducials (RGB) are present their co-ordinates (positions) are found else the control is transferred where again the camera continuously senses for colours.
- Now the obtained gesture (obtained from the knowledge of position of the colours) is checked for its validity i.e. if the obtained gesture is the pre-defined one. If the gesture is a valid it is recognized else again the control goes back to camera.
- Based on the recognized gesture, the following commands are decoded resulting in the following tasks:

• When the correct gesture is detected for the first time, the Microsoft PowerPoint presentation software opens.
• When the gesture is repeated for the second time, new slides are added.
• Similarly, when repeated for fourth time figure is inserted into the slide.
• And for the last time when repeated, the PowerPoint is saved and closed.

ii) Mode 3: Changing the slides in a PowerPoint

Third mode of operation is changing the slides of a PowerPoint.
The flow of the operation of mode1 is as follows:

- The camera here is a sensor and is programmed to continuously sense for the gesture and its detection.
- As soon as the image is captured and it is subjected to acquisition for further processing.
- After the image is acquired it is processed by performing different operations on it. The operations include implementation of different image processing algorithms. In our work, algorithms for colour detection (for example, Blob Algorithm) with better accuracy of detection and noise removal are implemented.
- After the implementation of the colour detection algorithm, the colours (RGB) if present are detected and the positions with intensities are obtained.
- If the fiducials (RGB) are present their co-ordinates (positions) are found else the control is transferred where again the camera continuously senses for colours.
- Now the obtained gesture (obtained from the knowledge of position of the colours) is checked for its validity i.e. if the obtained gesture is for the desired task. If the gesture is a valid one it is recognized else again the control goes back to camera.
- Based on the recognized gesture, the command that corresponds to changing of slides is decoded and the task is performed. This is performed continuously.

C) Summary of the Overall Operation

The summary of the overall operation can be described by a process called **PING-PONG** process.

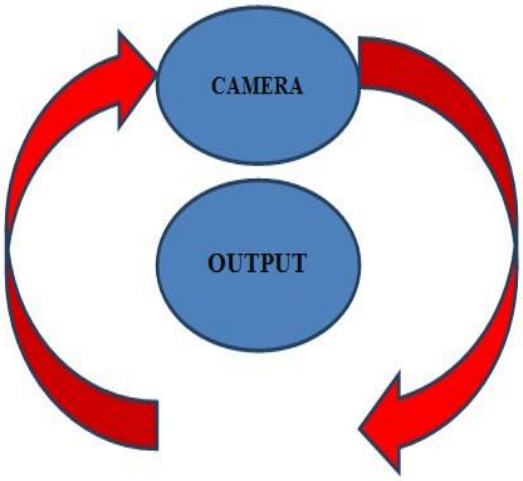


Fig 2: Ping-Pong Process

Camera: The camera that is acting as a sensor continuously senses[1]
for the gesture, captures it and sends for further processing. After sending the captured image camera waits for some time before beginning the sensing again.

Output: After the image[11] is received it is subjected to pre-processing, acquisition, application of algorithms, detecting for valid gestures, decoding the corresponding commands etc. and performing the intended tasks. After the task is completed and is displayed. The control is transferred to camera indicating that the system is ready for processing the next image. For the time, until camera sends another image to be processed, the output (system) waits by doing nothing.

4. ADVANTAGES AND DISADVANTAGES

A) ADVANTAGES

- **Cost effective**

This project incorporates low cost as it has minimal set of hardware.

- **Connectivity between real world and digital world**

Forming a connection between the real world and the digital world is the main aim of this project.

- **Direct accessing of data in real time [3][5][8]**

The user doesn't require any machine-human [2]interface to access the data. The data accessing is through recognition of hand gestures[4] and is much easier and user friendly compared to the text user interface or graphical user interface which requires keyboard or mouse.

- **Open source software**

The software that is used to interpret and analyze the data collected is an open source one. This enables other developers to contribute to the development of the system.

The other advantages are:

- This project is a user friendly interface which integrates digital information into the physical world and its objects, making the entire world your computer.
- This project does not change human habits but causes computer[17] and other machines to adapt to human needs.
- It uses hand gestures[4] to interact with digital information.
- It is gesture-controlled computing.

B) DISADVANTAGES

- Use of colour markers[9] might be undesirable to some of the users.
- User should have knowledge about the specific gesture for a specific task.
- Algorithm for colour detection should be chosen wisely.
- Laptop /PC is necessary as of now but can be eliminated in future.
- Increase in the distance of gesture from the camera can decrease accuracy of detection.

5. Simulation Results

MODE 1: Gesture 1

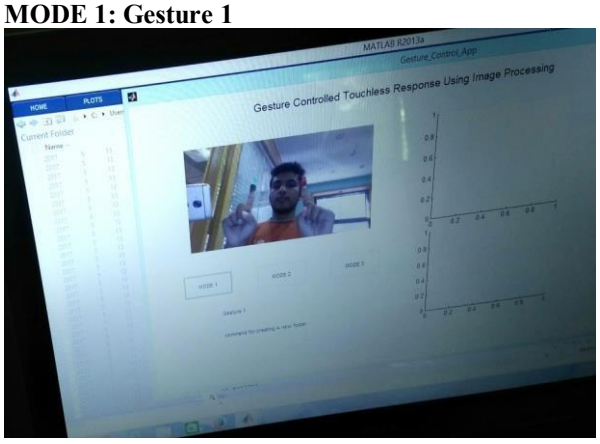


Fig 3: Image capture and save

MODE1: Gesture 2

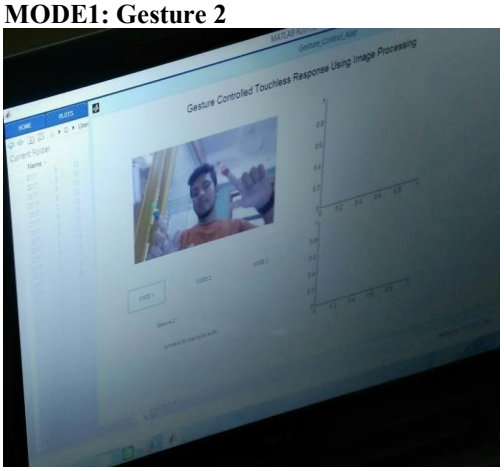


Fig4: To play an audio

MODE 1: Gesture 3

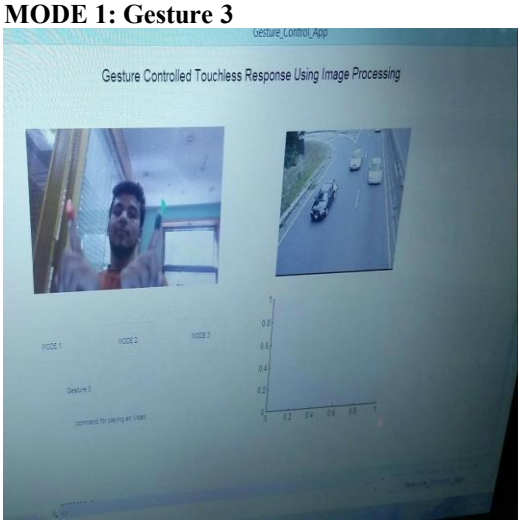
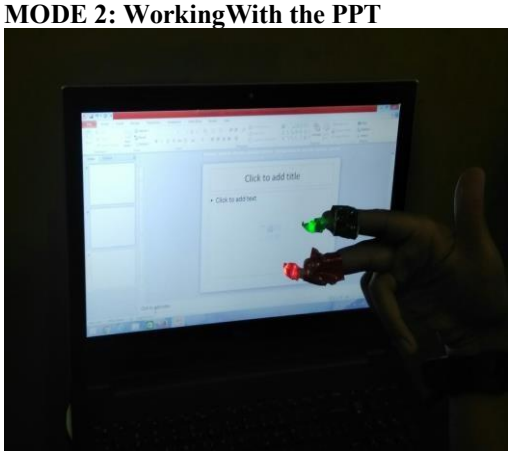


Fig5: To play a video

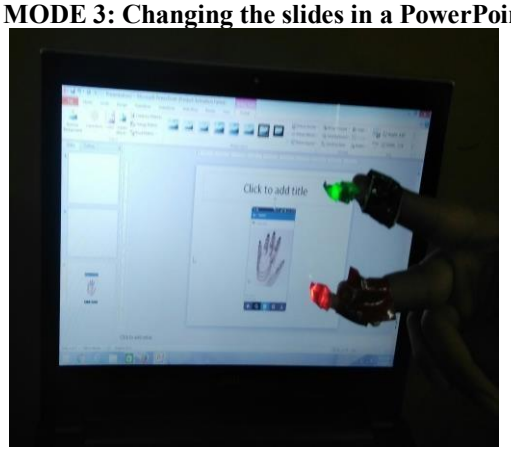
MODE 1: Gesture 4

Fig6: Image zoom-in and out

MODE 2: WorkingWith the PPT



MODE 3: Changing the slides in a PowerPoint



CONCLUSION

Our project was intended to perform various tasks on computer and we successfully performed tasks such as capturing an image, playing an audio, opening MS PowerPoint and working with it etc. with the help of gesture recognition and image processing principles. This is done only with the help of coloured sensors (caps/LEDs) that are worn on the fingers of the user. Our project is a gestural interface[6] that augments the physical world around us with digital information and lets us use our natural hand gestures[4]with coloured sensors to interact with that information. Our project enables humans to interact [2] with computers in a more direct way without using any external interfacing devices. Our project provides a much better alternative compared to the text user interfaces and graphical user interface which are usually used to interact with the computer.

REFERENCES

[1] M. K. Saha and S. Hore, "Sixth Sense Technology: A Brief Literary Survey,"Proceedings of International Journal of Engineering Research Technology (IJERT), Vol. 2 Issue 12, December 2013.

[2] V. J. Oniga, S. and I. Orha, "Intelligent human-machine interface using hand gestures recognition," In the proceedings of the IEEE Trans. On Automation Quality and Testing Robotics (AQTR), pp. 559 - 563, 24-27 May, 2012.

[3] I.-L. Jung, N. Akatyev, and Won-Dong Jang, "Touch less user interface based on marker detection and tracking for real-time mobile applications," In the proceedings of the International Journal of Innovative Computing(ICIC), Information and Control, Vol. 9 Issue 2, February- 2013.

[4] R. Z. Khan and N. A. Ibraheem, "Hand gesture recognition: A literature review," In the proceedings of International Journal of Artificial Intelligence Applications (IJAIA), Vol.3,No.4, July 2012.

[5] H. Park, "A Method for Controlling Mouse Movement Using a Real –Time Camera,"2008.

[6] Robertson P., Laddaga R., Van Kleek M., "Virtual mouse vision based interface", In the Proceedings of the nineth U international conference on intelligent user interfaces, pp. 177 U 183, January 2004.

[7] H. D. Nasser, "Hand gesture interaction with a 3d virtual environment," The Research Bulletin of Jorden, ACM , ISSN: 2078 U 7952 , Vol II(III), Page - 86.

[8] T. Kirishima, K. Sato, and K. chihara, "Real -Time Gesture Recognition by Learning Selective Control of Visual Interest Point," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 27, No. 3, pp. 351- 364, March 2005.

[9] M. Fiala and C. Shu, "3d Model creation using self-identifying markers and sift keypoints," In the proceedings of the IEEE Trans. on Haptic Audio Visual Environments and their Applications, pp. 118 - 123, 2006.

[10] Murthy and G.R.S., "Hand gesture recognition using neural networks", In the Proceedings of IEEE Trans. on Advance Computing Conference(IACC), pp. 134 - 138, 19-20 February,2010.

[11] Yubing Dong, Mingjing Li and Jie Li., "Image retrieval based on improved Canny edge detection algorithm", In the Proceedings of the International conference on Mechatronic Sciences, Electric Engineering and Computer (MEC), pp. 1453 - 1457, 2013.

[12] Ghosh D.K. and Ari S., "A static hand gesture recognition algorithm using k-mean based radial basis function neural network", In the Proceedings of 8th IEEE Trans. On Information, Communications and Signal Processing (ICICS), pp. 1-5, 2011.

[13] Kumar, S.P. and Pandithurai, O., "Sixth sense technology", In the Proceedings of IEEE Trans. on Information Communication and Embedded Systems (ICICES), pp. 947 - 953,2013.

[14] Bhowmick S., Kumar S. and Kumar A., "Hand gesture recognition of English alphabets using artificial neural network", In the Proceedings of IEEE Trans. on Recent Trends in Information Systems (ReTIS), pp. 405 - 410, 2015.

[15] Hsien-I Lin, Ming-Hsiang Hsu and Wei-Kai Chen, "Human hand gesture recognition using a convolution neural network", In the Proceedings of IEEE Trans. on Automation Science and Engineering (CASE), pp. 1038 - 1043, 2014.

[16] Chang Tan, Xiao and Nanfeng, "Improved RCE neural network and its application in human-robot interaction based on hand gesture recognition", In the Proceedings of 2nd IEEE Trans. on Information Science and Engineering (ICISE), pp. 1260 - 1263, 2010.

[17] Jalab, H.A. and Omer, H.K., "Human computer interface using hand gesture recognition based on neural network", In the Proceedings of 5th National Symposium on Information Technology :Towards New Smart World(NSITNSW),pp. 1-6,20