Manipulating Digital Information through Physical Manipulations

Arjun Toshniwal[1] Ankit Raghuram[2] Kishore Kalastri[3] Sharad Shriyan[4]
[1,2,3,4]Student
[1,2,3,4]Department of Electronics & Telecommunication Engineering
[1,2,3,4]SIES-GST, Maharashtra, India

Abstract— It is essential to link the digital world with the physical world so that the communication gap between the technologies is reduced to a considerable extent[2]. Thus the people who are new to this and doesn't have a fair amount of knowledge about coding and how the digital world operates will also be able to use it effectively .To ease our daily life with the technology around us it is necessary that technology understand our surrounding situations and help us with the required information. In order to achieve this desired goal between the technology and human it is need that technology works as our third eye [1].

Key words: IDE, Mouse Control, Real Time Text Extraction

I. INTRODUCTION

In present-day society, with the development of computer science, the software and hardware are very advanced. But the form of interaction between people and machine is developing slowly and even become an obstacle to the inherent technological development[8]. Therefore, a great number of researchers started to divert their attention from computer field of HCI has raised much attention and interest[5]. Many new computational applications have appeared and play increasingly important role in our life, such as tangible, robotics and gesture-based interaction.

Along with this we also wanted that computers should act as our third eye, i.e. the device should be able to understand the surrounding environment and process the input and give the desired output[7]. So as to achieve this desired target we are using the image processing techniques.

The project we have developed helps to ease the daily use of technology in our lives. It also helps disabled and uneducated people by providing a platform for using this technology in same way as others. Therefore, in the field of human computer interaction based on user-centered theory, gesture-based interaction has received great attention around the world and even is considered as the trend of future[3]. This paper basically studies gesture-based interaction.

It is motivated by the interest of a successful commercialized products, e.g xbox kinect,wii u etc. Compared with keyboard and mouse, gesture-based interaction is a more advanced technology and more natural form of human computer interaction (Dourish, P., 2004). However, except of its success in game industry, we haven't seen such good performance in other fields. Therefore, our research questions are what are the advantages and disadvantages of the gesture-based interaction? And from those what implications can we find for its practical usage in the future.

This stage of technology is just the beginning . In the coming days this combined idea of gesture command, image processing and text extraction can be developed to an advance level [6].

II. METHODOLOGY

A. *Modelling/Analysis:*

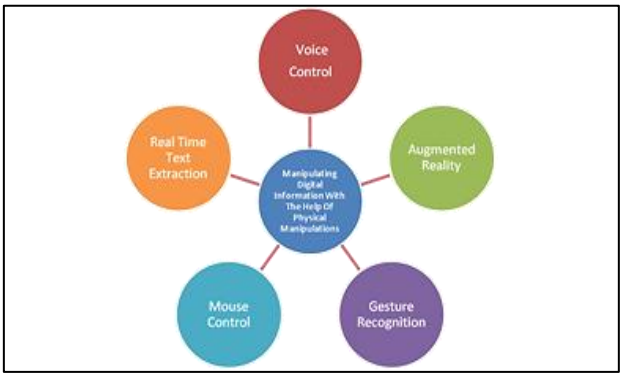

Fig. 1: Modelling / Analysis

The above block diagram represents basic idea of this topic. It consists of various modules each with their core functionality. These modules are then integrated into a single application .This allows the user to access each module through a single source.

B. *Software Implementation:*

The primary software used to create this application is Visual Studio 2010. Microsoft Visual Studio is an integrated development environment (IDE) from Microsoft. It is used to develop computer programs for Microsoft Windows, as well as web sites, web applications and web services. In this project Microsoft Visual Studio 2010 is configured with various other computer vision libraries. Once the configuration is over, further image processing is achieved with C++ and C# coding.

Since image processing is the basis of the project, OpenCV is preferred choice of computer vision library. The reason being it is open source and the library is written in C and C++ and runs under Linux, Windows and Mac OS. One of OpenCV's goals is to provide a simple-to-use computer vision infrastructure that helps people build fairly sophisticated vision applications quickly.

But since there are different modules written in different languages (viz. C++, C#) , a framework is required so that each language can use code written in other languages. Therefore .Net framework is used for this purpose. But to call OpenCV functions from .Net compatible languages such as C#, VB, VC++, IronPython etc. a .Net wrapper called as Emgu CV is used. The wrapper can be compiled by Visual Studio, Xamarin Studio and Unity, it can run on Windows, Linux, Mac OS X, iOS, Android and Windows Phone.

C. *Description of Various Modules:*

1) *Voice Control:*

This module deals with controlling computers through human voice.On giving certain voice commands the

computer is prompted to perform some action like opening facebook link, asking for current time etc. Here to recognise human voice Microsoft Speech API is used where the PC is first trained to recognise to human voice so that accurate results are obtained. Thus this part manipulates digital data through voice.

2) Augmented Reality:

This module deals with augmenting digital data on physical surfaces like wall, paper etc. To achieve this a computer tracking library is used known as ARToolkit. It is used to overlay virtual data on the real world. Here the application first identifies the unique marker and than uses that marker to augment digital data on it. For eg. A chess board is used as a marker, so as soon as the edges of chess board are detected a video can be played on the chess board. Thus in this module digital data is manipulated in real world.

3) Gesture Recognition:

This module deals with recognizing real time hand gestures. Each gesture can be stored in the database and recognized accurately. The accuracy of recognition depends on the training of hand gestures. Haarcascades training (haar training) is a quick tool to achieve accurate hand gesture detection and recognition[4]. We need to collect positive images that contain only objects of interest, e.g., hand. After that collect negative images that does not contain the object of interest. The number of negative images should be twice or thrice that of positive images. This will produce more accurate results. So once positive and negative samples are created these samples are trained which might take days depending on the number of samples. As a result of training a .xml file is created which can be used to recognize gestures. Thus unique gestures can be recognized with the help of haar training and subsequently it can be used to create sign language.

4) Mouse Control:

This module deals with controlling mouse with the help of hand. Thus human hand can act as a physical mouse which controls the computer. For this purpose cvBlob library is used. cvBlob is a library for computer vision to detect connected regions in binary digital images. cvBlob performs connected component analysis (also known as labeling) and features extraction. Thus using this library coloured markers are uniquely identified and thus on identification respective mouse controls can be assigned to specific colour or specific colour set.

5) Real Time Text Extraction

This module deals with real time extraction of text as captured from webcam. To uniquely identify text Tesseract engine is used . Tesseract is an optical character recognition engine for various operating systems. It is a free software, released under the Apache License, Version 2.0. Region of interest where the text is present is identified first , than using tesseract engine the text is extracted from the image and identified accordingly. This engine does not recognise special characters and deals with pure text. Thus using this engine text in various languages can be identified.

III. Conclusion

As the world moves towards cloud computing and advanced image processing one can only appreciate the benefits of this technology and the bright future it beholds.This a improved human-computer interface not only allows the regular user to achieve machine control and machine learning without external hardware but also allows the disabled and the handicapped to access the same..In India where traditional computers are being replaced by smartphones and tablets,this technology takes this accessibility even further, not to mention the benefits it provides to many industries and institutions

IV. Acknowledgement

We like to offer sincere thanks to our project guide Prof. Tejal Tirodkar who has helped us throughout the compilation of our work.

Also we would like to express our sincere gratitude to our Principal Prof. Alka Mahajan for giving us the opportunity to enhance our knowledge and for allowing us to carry our work on the project.

We also like to thank all the other staff members and non-teaching staff for their support and help.

References

[1] P. Premaratne, "Historical Development of Hand Gesture Recognition-, Human Computer Interaction Using Hand Gestures", Cognitive Science and Technology,© Springer Science+Business Media Singapore 2014

[2] Prof. Mr. D.S. Patil1, Mr. Shahak Patil, "Sixth Sense Technology-A New Innovation ",International Journal on Recent and Innovation Trends in Computing and Communication

[3] Swarali Narvekar and Manali Godse, "Vision based Analysis using Sixth Sense Technology" for International Conference and Workshop on Emerging Trends in Technology 2013

[4] P. Viola and M. Jones. "Rapid object detection using a boosted cascade of simple features." Computer Vision and Pattern Recognition, IEEE Computer Society Conference on, 1:511, 2001.

[5] Xu Yan ,Nuerrennisahan Aimaiti, "Gesture-based interaction and implication for the future" ,UMEA universitet,2011

[6] Vladimir I. Pavlovic, ,Rajeev Sharma, Thomas S. Huang, "Visual Interpretation of Hand Gestures for Human-Computer Interaction: A Review", IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE, VOL. 19, NO. 7, JULY 1997

[7] Harsha vaccher," Need for Gesture Recognition"Engineering and R&D dept,HCL, April 2014

[8] Monika arora, "Basic Principles of Sixth Sense Technology" VSRD-IJCSIT, Vol. 2 (8), 2012, 687-693

