INDIAN INSTITUTE OF TECHNOLOGY KANPUR

PROJECT REPORT

Kinect Hands Free

Submitted by
Agrim Bari
Roll No. 150057
Department of Electrical Engineering
Indian Institute of Technology
Kanpur – 208016
Email id: agrimb@iitk.ac.in

Under Guidance of
Dr. Soumya Ranjan Sahoo
Department of Electrical Engineering
Indian Institute of Technology
Kanpur – 208016
Phone No: 0512-2597944
Email id.: srsahoo@iitk.ac.in

CERTIFICATE

This is to certify that the project entitled "**Kinect Hands Free**" submitted by Agrim Bari (150057) as part of Summer internship offered by the Indian Institute of Technology, Kanpur, is a bonafide record of the work done by him under my guidance and supervision at the Indian Institute of Technology, Kanpur from 15th May ,2016 to 11th July ,2016

Dr. Soumya Ranjan Sahoo

TABLE OF CONTENTS

1.Abstract

This project describes the design and implementation of the speech and gesture recognition system to control inputs in the computer using the Microsoft Kinect.This system focuses on the identification of natural gestures that occur during computer and user interface, making the user experience as fluid as possible. The system uses an HMM to classify the performed gestures in conjunction with an SVM to perform real-time segmentation of gestures. The fusion of these two models allows the system to classify gestures in real time as they are being performed instead of waiting until completion. The incorporation of speech commands gives the user an additional level of precision and control over the system.

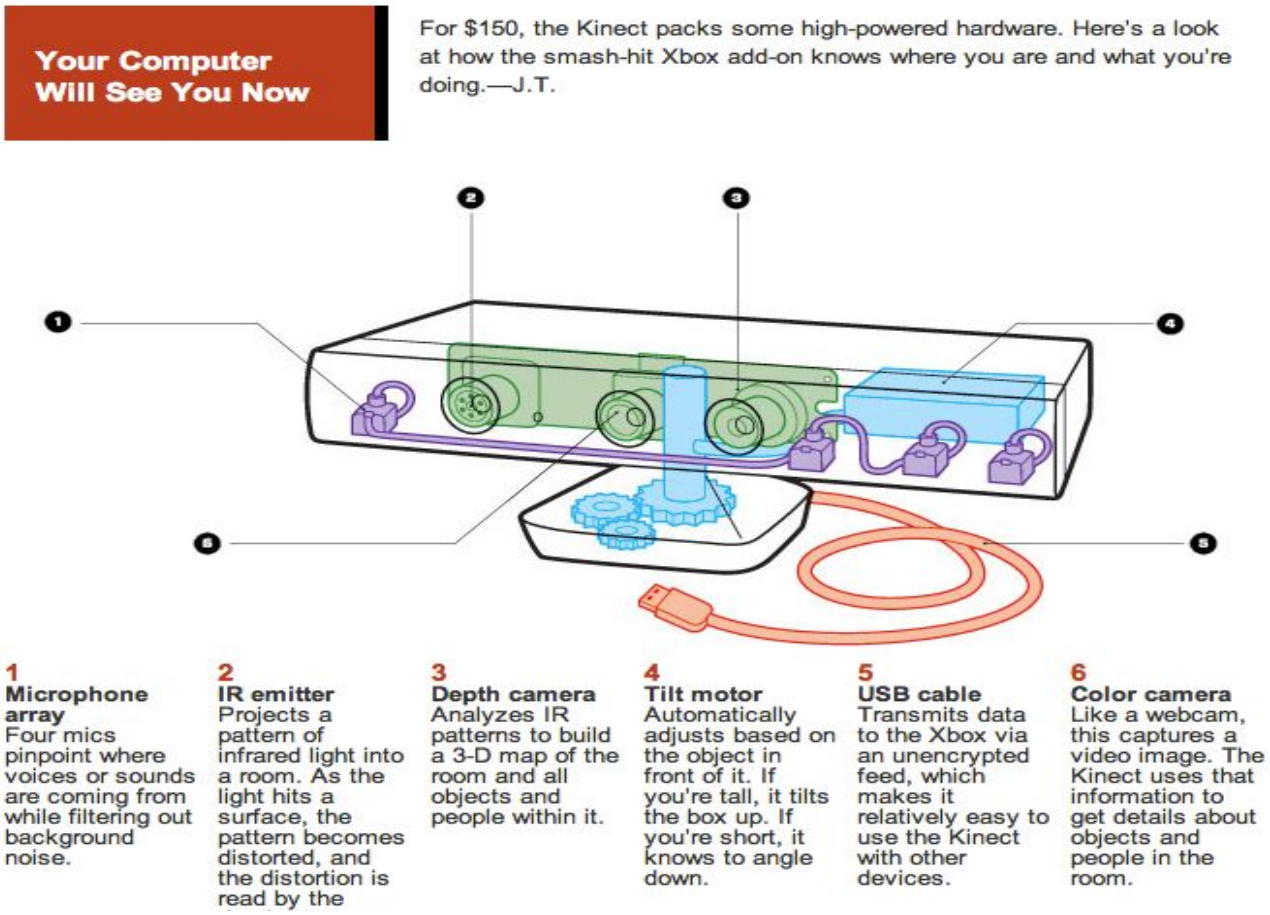
2.Introduction

2.1 Motivation

In today's world, technology pervades nearly every aspect of the average person's daily life. People interact with computers and other technology as frequently as they do with other people and they should have the ability to communicate with computers as naturally as they do with other humans. Speech is perhaps the most comfortable form of communication between humans. It is quick, efficient, and allows people to express themselves with great degrees of freedom, limited only by their own vocabulary. Since the dawn of computers half a century ago, people have dreamed of being able to have conversations with robots and other artificial intelligences as easily as they do with other humans. Unfortunately, keyboard and mouse have been the primary means of interfacing with computers even to this day. While effective in many situations, they are limiting and not a particularly natural means of interaction. Gesture recognition is a current area of research that is trying to address this problem. Everyone is familiar with gestural interaction with other humans. It occurs naturally during speech as a way people for people to express themselves. Gestures are a form of body language that are essential to effectively communicate ideas in addition to spoken language. People already gesture when communicating with other humans, so why not use this mode of communication for natural interaction with computers.

3. Working

The Kinect performs 3D depth sensing by emitting a structured point cloud pattern of infrared (IR) light and calculating the depth from the images taken with its IR sensor. Because this IR point cloud originates at a single point in the Kinect sensor, as the distance from the Kinect increases, the point cloud pattern disperses proportionally to the distance the light has traveled. By measuring the offset between the expected location of the IR grid pattern at a calibrated distance and its actual location, the Kinect can calculate the depth at each point in the projected point cloud. Using this depth image, the Kinect can identify foreground objects and determine people and their poses by comparing the detected body to millions of stored examples of body poses. The Kinect then uses a randomized decision forest technique to map the body depth image to body parts from which the skeleton representation can be built .



4. Implementation

4.1 Hardware and Software

1. Microsoft Kinect 1.0 is the main hardware component of our project, developed by Microsoft. The Kinect is a very versatile equipment and has been put to use in various spheres of life. It has various inbuilt sensors such as -

1. An RGB Camera
2. An infrared sensor

3 Microphone Array

In our project we are mainly making use of three of these sensors. Using the
Infrared sensor and RGB Camera data we are processing the body movements(Gestures Recognition).

2. Kinect for Windows SDK v1.8
The Kinect for Windows SDK provides the tools and APIs(Application Programming interface), both native and managed, that you need to develop Kinect-enabled applications for Microsoft Windows

3.Microsoft Visual Studio 2015

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 our project we made use of these softwares systems, which provides us the set of drivers to get data from the Kinect and thereafter code it to use it our own way.

5. Limitations and Future Scope

The data sent by Kinect 1.0 is not accurate enough to be able to process Facial details and finger movements as a result we have not been able to implement Facial Recognition or Finger movements.The project can be continued further in the future using the Kinect v2.0 to include Facial Recognition and delicate finger gestures and further smoothness. Thus a fully interactive user friendly environment can be done.

6. Links and References

6.1 Github

- **https://github.com/Pranjalgiri/kinect-handsFree**

6.2 Readings

- http://users.dickinson.edu/~jmac/selected-talks/kinect.pdf
- http://pterneas.com/2014/01/27/implementing-kinect-gestures/
- https://www.youtube.com/watch?v=uq9SEJxZiUg
- https://github.com/FTSRG/publication-pages/wiki/Realtime-gesture-recognition-with-Jnect-and-Esper
- https://github.com/Marsyangkang/KinectPowerpointControl/blob/master/KinectPowerPointControl.sln
- https://dspace.mit.edu/bitstream/handle/1721.1/85410/870310033-MIT.pdf?sequence=2