# Introduction:

Problem and importance?

1. Gesture recognition
   1. Natural interaction
   2. Interpret human gestures and use that as input to a system.
   3. Example move cursor according to hand movement. Or convert sign language to human speech.
2. Input Devices:
   1. Kinect
   2. Wired gloves
   3. Leap motion Controller
   4. Myo
3. Leap Motion

# Background:

Literature survey

1. All the current work

# Solution

What is being built and how the system will address the problem.

# System Implementation:

Details

Milestones

How to evaluate

# Question:

Gesture Recognition is a field that focuses on understanding the human body language and interpreting them so that it could be used as input to system. For example we use Kinect as input to certain gaming systems. It is relatively an old topic but there is a lot of development works recently with primary focus on using hands and face gestures as input. This would change the way how we interact with the system and could replace the traditional keyboard and mouse input.

There is a large set of input devices available for gesture recognition but they could be categorized in to two major categories. Vision based and non-vision based. Vision based sensors make use of camera to capture the human gestures devices include Kinect, leap motion and Intel real sense cameras. Non vision based devices uses accelerometer data, data from ECG sensors like Myo which reads the muscle movements and interpret them.

My work will be based on leap motion so let’s see a little about it. Leap motion Controller is publicly available from May 2013 with its second version from May 2014. It has two IR cameras and three LEDs and it uses it to track the hand and finger gestures accurately in real time. It records data at the rate of 50 – 200 frames per second. It’s highly accurate to the range of 0.01mm but it has limited observational area. It’s the cheapest device available in the market and it makes natural interaction possible to any computer system as it works with all major operating systems.

My work in this project will mainly focus on implementation of various gesture recognition algorithm and build a library for Leap motion sdk so that the gesture recognition task would be easier. My initial thoughts is to build a verbose library so that it could recognize many gestures and tie it to any actions available in the sdk. The library will be built using java and will be made publicly available. The subtasks of the project is to collect data from the library and use it to compare various algorithm in terms of its effectiveness and responsiveness and provide any improvement if possible.

That been said let’s look in detail on what is going to be implemented. Human gestures can be classified into two major categories called action and parametrized gestures. Action gestures include two subclasses static and dynamic gestures. Static gestures are those where only one action is involved and it does not change over time. Like the stop sign while dynamic gestures change over time like hand movement to move mouse pointer or scroll action. Parameterized gestures have some sort of parameter associated with them like pinch and zoom where amount of zoom depends on the distance between the fingers.

Let’s look into some algorithms that is used for gesture recognition. Various methods were proposed to model the gesture recognition problem. Some of them include HMM, NN, SVM, SPRM, FSM, and TDNN. These are most popular ones and the research papers talks about some modifications or variations of these algorithms. The accuracy of these algorithm varies from 80 to 99.6%.

HMM is one of the first proposed and most commonly used algorithm in many recognition and pattern matching field. It’s a statistical model based on Markov’s property which states that the future state of any process depends only on the present state and the probalities for transition between the current and future states. In HMM the output is clearly known but not the states involved. For example if u pinch and zoom the output is clearly known but not the states involved as it might denote another gesture like click of the mouse. The example at the side is from Wikipedia where a doctor is about to predict whether the patient is healthy or has fever. But he doesn’t know them directly as its hidden from him. All he knows is how the patient feel and use the Markov’s property to calculate it. The property and model has a wide range of use in many fields and is mostly accurate as it’s based on conditional probability. Various algorithms like forward backward algorithm, Viterbi algorithm were used to tackle HMM programmatically.

Another major algorithm is NN. NN are modeled based on human brain and has three layers of neurons input output and hidden. Each neuron has a weigh associated with it and these weights are self-adjusted during the initial training process. If trained properly NN will produce the most accurate results. But it’s impossible to train them to include all human gestures plus sometimes it’s computationally expensive. Once training is complete the activation functions are used to determine the output of the gestures.

Mile stones