

**Military Technical College
Kobry El-Kobbah,
Cairo, Egypt**



**7th International Conference
on Electrical Engineering
ICEENG 2010**

Hand Gesture Recognition for Human-Computer Interaction Through Infrared Motion Sensing

By

Jan Adrian Adriano*
Jay K Delicana*

Enrico Aquino*
Karl Jamoralin*
Cyrel Ontimare*

Christopher Cabael*
Noel Linsangan*

Abstract:

This paper presents a method of computer interaction using the IS-300Pro Precision Motion Tracker. The motion tracker compares the gestures with a defined pattern of movement so that it can be recognized. The method is developed using infrared light emitting diodes (IR LEDs) attached to the thumb and index fingers of two gloves, and the IS-300Pro motion tracker connected to a computer using the USB port. A microcontroller unit (MCU) supporting USB connection is employed for computer interface. Software developed specifically for the application is installed in the computer. The system is able to recognize gestures for alphanumeric characters, as well as user-defined gestures. The software interprets the gestures captured by the motion tracker and translates them to corresponding response hardcoded in the software. Three software programs were developed to demonstrate the capabilities of the system. Experiments show that distance between the sensor and IR LEDs affects the accuracy of gesture recognition and interpretation.

Keywords:

Motion Tracker, Gestures, Infrared, Hardcoded, Gesture Recognition

*Mapua Institute of Technology, Manila, Philippines

1. Introduction:

[1]A gesture is a form of non-verbal communication made through a part of the body; most common are the hands which are very expressive in terms of communication. The language of gesture allows individuals to express a variety of feelings and thoughts, from contempt and hostility to approval and affection. Hand gestures are performed by one or two hands including the fingers, it is the most numerous categories of gestures due to the ability of the human hand to acquire a huge number of clearly discernible configurations.

The research hypothesized that if gestures can be detected and interpreted correctly, then it can be used to influence the operation of a computer system. Specifically, to simulate basic computer operations using finger gestures.

[1]The conventional personal computer is controlled with at least two input devices, the mouse and the keyboard. Computer users are accustomed in using the PC with these two input devices. However, these input devices can only be utilized efficiently when the user is in a stationary position, such as a sitting position. This restricts the user interaction with the computer in a short distance. The proponents decided to create a device that can be an alternative way of interacting with the computer. This device will give the user capability to interact with the computer from a certain distance more than that with the mouse and keyboard.

[1,3,7,8,9,10]Contemporary research into gesture recognition focuses primarily on raising the level of interaction between humans and computer system. [1,7,10]Gesture recognition is useful for processing information from humans which is not conveyed through speech or type. Current studies use gesture recognition primarily to aid persons with disabilities by means of recognizing movements corresponding to commands for a certain system.

[1]The ability to track a person's movements and determine what gestures they may be performing can be achieved through various tools. [2,4]Although there is a large amount of research done in image/video based gesture recognition, there is some variation within the tools and environments used between implementations.

[1]Gesture recognition is a topic in computer science and language technology with the goal of interpreting human gestures via mathematical algorithms. [1,10,12]Gestures can originate from any bodily motion or state, but commonly from the face or hands. Many approaches have been made using cameras and computer vision algorithms to interpret sign language. However, the identification and recognition of posture, gait, proxemics, and human behaviors is also the subject of gesture recognition techniques. It can be seen as a way to convey information from humans other than speech or type.

[1,2]Motion can be detected by measuring change in speed or vector of an object or objects in the field of view. [2]This can be achieved either by mechanical devices that physically interact with the field or by electronic devices that quantifies and measures changes in the given environment. When motion detection is accomplished by natural organisms, it is called motion perception.

[2,3,5,7]Principal methods by which motion can be electronically identified are optical detection and acoustical detection. Infrared light or laser technology may be used for optical detection. [2,3]Motion detection devices, such as motion detectors, have sensors that detect movement and send a signal to a sound device that produces an alarm or switch on an image recording device. [2]There are motion detectors which employ cameras connected to a computer which stores and manages captured images to be viewed later or viewed over a computer network. Figure 1 shows an example of a Passive Infrared Detector (PID). A PID is an electronic device used to detect infrared light emitting from objects in its field of view. Figure 2 shows a Pyroelectric sensor used to detect infrared light. The pyroelectric sensor is made of a crystalline material that generates a surface electric charge when exposed to heat in the form of infrared radiation. When the amount of radiation striking the crystal changes, the amount of charge also changes and can then be measured with a sensitive FET device built into the sensor. The sensor elements are sensitive to radiation over a wide range so a filter window is added to the TO5 package to limit detectable radiation to the 8 to 14mm range which is most sensitive to human body radiation.



Figure 1: Commercial PID w/ cover removed showing the location of the

Pyroelectric sensor

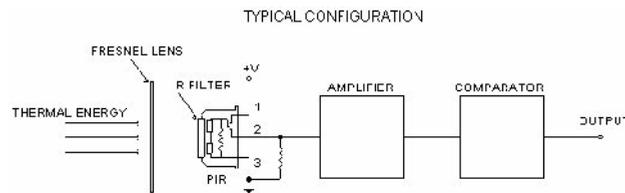


Figure 2: FET configured as a Pyroelectric sensor

Statement of the Problem

Mainstream natural human-computer interaction such as gesture recognition application is not available to the public. Almost all of them are used only for research and study due to the cost of building devices for commercial use. Most information from humans cannot be conveyed through speech or type. Currently, only a few systems exist that directly influences a computer system by means of hand gestures. And still an even fewer devices are brand-specific, making it hard to customize and be used with other applications.

Objectives of the Study

This study aims to present a method of computer interaction through the use of finger gestures. The gestures corresponding to characters or commands can be customized depending on user's preferences. The study also aims to present a low-cost human-computer interface.

Scope and Limitations

This present study covers the following topics only.

1. The design can be used to make use of finger movements to control and manipulate software applications.
2. The design can perform basic commands such as zoom in, zoom out, rotate, scroll, and opening programs.
3. The software system recognizes strokes corresponding to alphanumeric characters
4. For it to be recognized, the action should be single-stroke, meaning the pinch should not be released to complete the stroke.

At this stage of its development, the study has the following limitations

1. Line-of-sight contact between infrared (IR) light-emitting diode (LED) and IR sensor.
2. Use of only two fingers, the thumb and the index finger of each hand
3. Inability to recognize multi stroke characters.
4. The use of the “Cybergloves”, devices fitted with infrared (IR) light-emitting diode acting as the source.
5. The system should be operated in a low light environment to minimize the infrared radiation detected by the sensor.
6. The Character Recognition program cannot distinguish between capital and small letters.

2.Literature Review:

Several studies have been made on alternative input devices or systems as well as related technology.

[1]The core of our present work was the study in 2004 of Jean-Christophe Lementec and Peter Bajcsy. They used so-called cyber gloves equipped with 18 IR emitters to show finger movement and finger articulation. The IR light were then detected and interpreted by an IS-300Pro Precision Motion Tracker manufactured by InterSense. The study focused on gesture classification, identifying every arm movement of the user and comparing low-level gesture characteristics in a defined pattern of movements for it to recognize the gestures. The IS-300 Precision Motion Tracker did not sense the transition of movements, from previous movement to a current movement but was able to recognize eleven gestures. Some of the valid gestures in this study were move ahead, turn left, pass control and slow down. In analyzing the study, it helped the researchers to know how the motion tracking device should function and how it behaved in identifying and classifying valid gestures.

[2]A neural network utilizing gesture recognition is used in order for the computer to identify user movements and determine if a movement is a predefined gesture. This technique is also used in the present study to establish any viable pattern for the motion of tracked dots. The combination is described by Grossberg.

[4]A similar study conducted in 2006 entitled “Automatic video surveillance using statistical analysis of temporal posture sequences” presents an idea of using video camera in order to capture and recognize human activities. The method is a three-step approach to recognize human activity. One is human activity segmentation using motion analysis. Then after segmentation, the body posture is estimated. Afterwards, a temporal model of detected posture is generated using hidden Markov models for activity identification. Particularly in the second step, the camera should detect segment points on the human body in order to identify the posture. This is one of the fundamental concepts of motion capture.

[6]The study entitled “Robust human posture analysis using incremental learning and recall based on degree of confidence of feature points” conducted in 2009 introduces an improved method of detecting feature points in the human body. The study uses a vision-based motion capture system (MCS) using the variable-density self-organizing map (VDSOM). Self-organizing maps falls on the category of neural networks, in which visualization of low-dimensional views of high-dimensional data is of priority. It aims to mimic how the human brain operates, thus relating to how a computer behaves using different input data.

[2,5]Currently, research into gesture recognition focuses on raising the level of human and machine interaction. Several studies were conducted demonstrating the trend in pervasive computing that focuses on natural interaction with a system. Pervasive or ubiquitous computing means that the system is small enough not to be noticed easily by the user.

[2]One study conducted in 2007 used an intelligent wheelchair that recognizes head and hand movements as input commands. The study entitled “Head gesture recognition for hands-free control of an intelligent wheelchair” utilizes visual recognition of head movement and hand gestures using a video camera. The study used Adaboost face detection algorithm and Camshift object tracking algorithm for movement detection and facial recognition.

[10]A study published in 2010 adapted from the Nintendo Wii controller uses a 3-axis accelerometer to detect hand gestures and recognize posture as a means of input in a computer system. The study entitled “High-level programming and control for industrial robotics: using a hand-held accelerometer-based input device for gesture and posture recognition” aims to teach industrial robots by mimicking the gestures of the hand on which the Wii controller is attached, then repeating that movement indefinitely. The interaction is more natural than using a teaching pendant used in programming

industrial robots. Tests results show that the system works even in non-controlled environment.

[5]The study conducted in 2009 entitled “Human-computer interaction in ubiquitous computing environments” explores the possibilities of using hand gestures as a means of interaction between humans and computers. The paper discusses the quantification of human movement using Fitt’s Law in order to model human-computer movement. The law predicts that the time required to rapidly move to a target area is a function of the distance and size of the target. The study also discusses some of the human factors and technical considerations that arise in trying to use the human body as a means of input.

[7]“Gesture-based human-robot interaction using a knowledge-based software platform” a study conducted in 2006 used a frame-based knowledge model to describe a gesture-based human-robot interaction system (HRI). Frame-based model uses the concept of inheritance, in which properties of a previous frame is forwarded to the next frame. Using this knowledge, the paper hard-coded a software platform that recognizes facial patterns and hand gestures. Tests results showed that the system is capable of recognizing static as well as dynamic hand gestures, though only static face poses. The paper provides an experimental HRI implemented in frame-based software platform.

Figure 3 shows the character strokes used by current personal digital assistant (PDA) devices.

Write lowercase letters on LEFT side, and capital letters across MIDDLE of input area							
Letter	Stroke	Letter	Stroke	Letter	Stroke	Letter	Stroke
A	Λ	B	β	C	c	D	d
E	ε	F	ƒ	G	g	H	h
I	i	J	j	K	k	L	l
M	m	N	n	O	o	P	p
Q	q	R	r	S	s	T	t
U	u	V	v	W	w	X	x
Y	y	Z	z				

Figure 3: Character stroke patterns on personal digital assistant

3. Methodology:

Infrared Motion Tracker Hardware

Figure 4 shows the infrared motion sensor. The motion tracking hardware consists of the IS-300Pro Precision Motion Tracker manufactured by Intersense. It identifies every arm movement of the user and compares low-level gesture characteristics in a defined pattern of movement for it to recognize gestures. Figure 5 shows the unit in which the motion tracker is connected that interfaces with a computer via USB port. Figure 6 shows the block diagram of the system hardware.



Figure 4: IS-300Pro Precision Motion Tracker

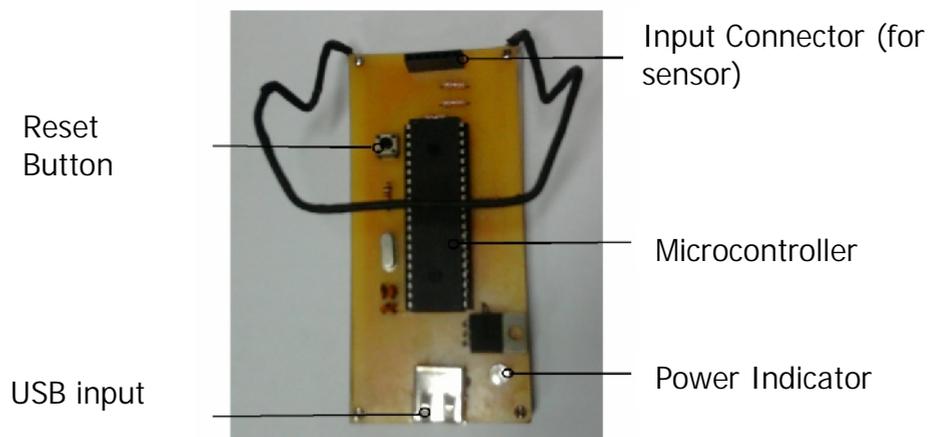


Figure 5: Base unit with MCU for PC interfacing

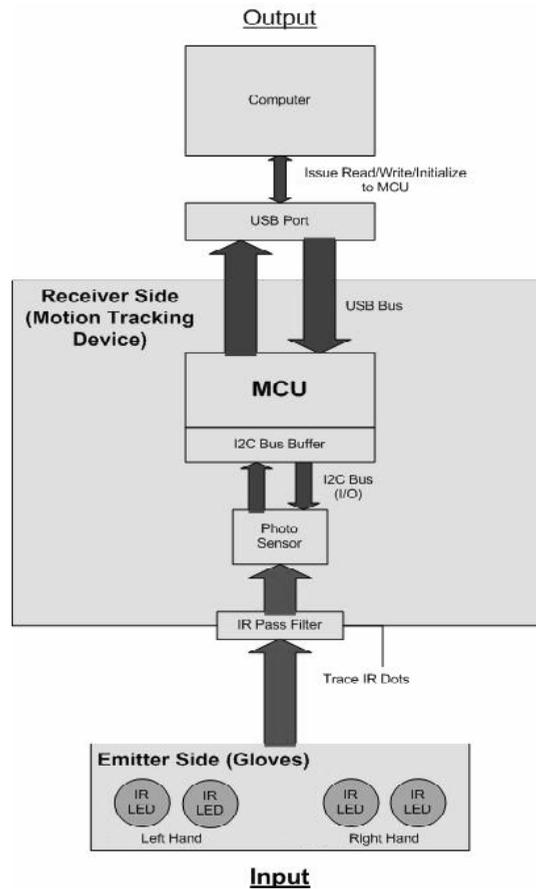


Figure 6: Prototype Block Diagram

Gesture recognition software

Figure 7 shows the gesture recognition software engine. The MCU stores 12 bytes of data obtained from the photosensor. Since the USB is only capable of sending 8 bytes of data at a time, a dynamic link library (DLL) was created to communicate with the microcontroller. The DLL sends a signal to the MCU after the first 8 bytes of data, to send the last 4 bytes, for a total of 12 bytes. The engine flowchart consists of three major phases, the initialization, refresh and check phase.

a. Initialization phase

The Initialization phase introduces all the globally declared variables and initializes the DirectX9 engine. DirectX9 mode allows the users to draw 3D mesh models and also allows the user to take advantage of the transformation matrix system available in DirectX9. The transformation matrix system allows objects to be easily

repositioned, scaled and rotated without complex mathematics. This makes it easier for the designers to program the codes.

b. Refresh phase

The second phase is the refresh stage where the application calls the DLL to return the 12-byte point properties including the X and Y coordinates of the 4 points. The refresh stage also redraws all the graphical elements on the screen like the dot representation of the infrared LEDs. After refreshing the values, the application checks if the user sent a “program close” command such as escape key to close the engine.

c. Check phase

The third phase is the check phase where the application checks whether a pinch has been detected. The program does this by scanning whether a point suddenly disappeared and comparing it with its closest point. If the distance between the two points is very small then a pinch is registered. It temporarily stores the object location to differentiate the movements of the dot. After that, the application returns to the refresh phase.

The gesture application uses a path comparator algorithm to differentiate between user strokes and stored strokes. Figure 8 shows how the initial dot track is transformed into an array of data that the computer can understand. When the application recognizes a pinch, it starts to record the dot x and y per cycle. The locust of X and Y are divided into 50 (median percentage) individual segments of dots. Finally, the 50 dots are converted into vector angles. The 50 angles are cross checked with the predefined strokes (each with 50 angles) the stroke with the least error percentage is chosen as the winner and interpreted by the computer as the stroke of choice.

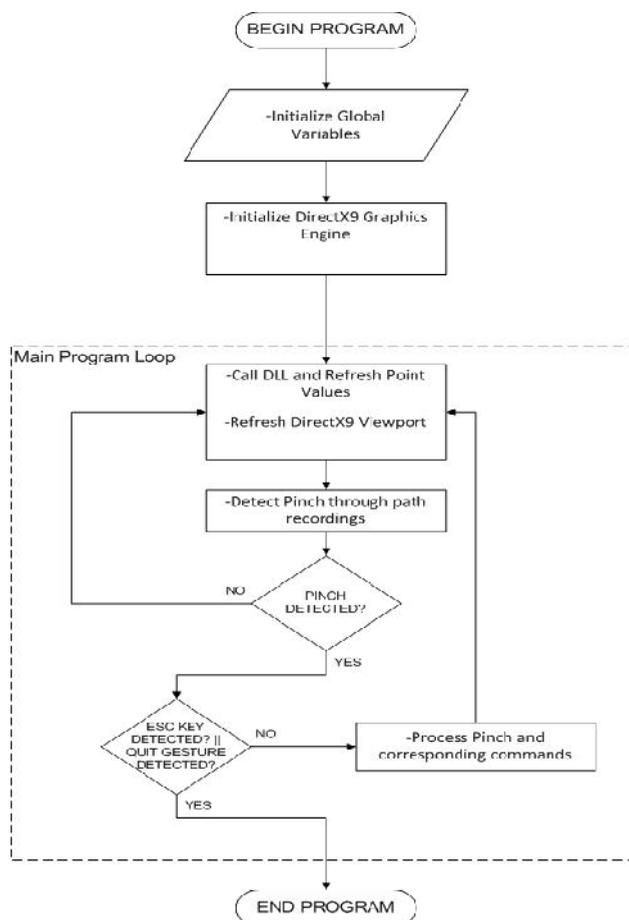


Figure 7: Software Engine Flowchart

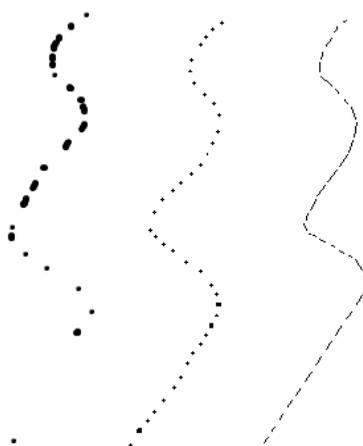


Figure 8: From pinch X and Y to Vector data

Infrared Motion Tracker Input device

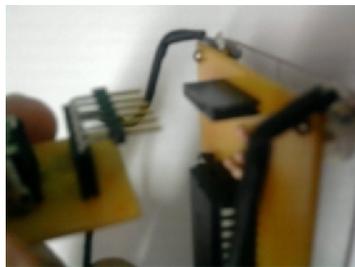
[1]Figure 9 shows the “Cybergloves”. A pair of gloves is used as an input device for gestures. Infrared light emitting diodes are placed on the thumb and index fingers of both gloves. The gloves are colored black to allow maximum detection of the IR LEDs. The gloves can be used singly or in a pair. Detection of gesture begins when a pinch is detected by the sensor. The gesture ends when the pinch is released.



Figure 9: Gloves fitted with infrared light emitting diode at the finger tips

Hardware installation

1. Connect the sensor to the unit using either the straight connector (for flat surface mounting) or the elbow connector (for laptop mounting).



2. Connect the USB cable to a USB port of the computer and the MCU circuit.



Operating Procedure

1. Remove the strap located in the wrist of the gloves.
2. Insert your hand in the corresponding gloves to be used.
3. Reattach the strap and make sure it is tight.
4. Insert the battery in the gloves.
5. Switch on the IR LED of the gloves by moving the switch upward.
6. In the computer click on the shortcut or go to the start menu then run the software application.

Software Applications Installed

1. Satellite Demo - loads an image to the screen in which the user can do functions such as zoom, rotate, and drag
2. Scrapbook Demo - loads multiple images to the screen from which the user can select a specific image and performs functions such as zoom, rotate and drag
3. Input-Gesture Demo - loads a blank screen that recognizes patterns for opening a specific application program
4. Character Recognition – recognizes user-defined gesture patterns corresponding to alphanumeric characters

Using Gesture Command

- By using the Satellite and Scrapbook Demo

Grab

- Position the dots in the picture and then make a pinching gesture to grab the picture.

Zoom

- To zoom in, grab the picture using both hands and then move your hands closer to each other. To zoom out, grab the picture using both hands and then move your hands farther to each other.

Rotate

- To rotate, grab the picture using both hands and then make a rotating gesture to rotate the picture.

- By using the Input Gesture Demo

Note that these gestures are defined by the user of the system.

Internet Explorer

- Make a pinch gesture, and then draw an “e” into the screen.

Microsoft Application

- Make a pinch gesture, and then draw a “P” into the screen.

Exit Input Gesture Demo

- Make a pinch gesture, and then draw a “ ” into the screen.

Safety Precautions

1. Do not touch the battery while in use.
2. Dispose used battery properly
3. Do not touch any electronic components of the device.
4. Avoid exposing the batteries and the unit to high temperature.
5. Store the device in a safe place.

- Character Recognition program

The software contains user defined gesture patterns representing alphanumeric characters. To generate the character, the user needs to gesture the predefined pattern stored in this application software.

4. Experimentation and Analysis:

Experimentation

Three applications were developed in order to demonstrate the capabilities of the system, namely, Satellite demo, Scrapbook demo and Input gesture demo. The satellite and scrapbook demo are similar applications that demonstrates basic computer commands such as zoom, rotate and scroll. Input gesture demo uses gesture recognition to issue commands to the computer. These commands are hardcoded into the system software. Table 1 shows the results of the satellite and scrapbook demo test. Several gestures were performed at incrementing distance from the sensor. R stands for recognized, NR for not recognized. For pinch to be detected, the two dots on the screen should merge, meaning the gesture was recognized. After pinch was detected, the dots should move around as the user moves the gloves indicating a successful drag.

Table 1: Satellite and Scrapbook demo showing gesture detection at different distances from the sensor

Distance (cm)	Trials	1 Pinch	Pinch + Drag	2 Pinch	2 pinch + Drag
50	1	R	R	R	R
50	2	R	R	R	R
50	3	R	R	R	R
100	1	R	R	R	R
100	2	R	R	R	R
100	3	R	R	R	R
150	1	R	R	R	R
150	2	R	R	R	R
150	3	R	R	R	R
200	1	R	R	R	R
200	2	R	R	R	R
200	3	R	R	R	R
250	1	R	R	R	R
250	2	R	R	R	NR
250	3	R	R	NR	NR
300	1	NR	NR	NR	NR
300	2	NR	NR	NR	NR
300	3	NR	NR	NR	NR

Table 2 shows the results of the pattern recognition test. The pattern used in this test is the infinity symbol. The gesture is performed at increasing distance from the sensor. R stands for recognized, NR for not recognized.

Table 2: Pattern Recognition test showing gesture detection at different distances from the sensor

Distance (cm)	Pattern Recognition (Infinity Symbol)
50	R
50	R
50	R
100	R
100	R
100	R
150	R
150	R
150	R
200	NR
200	NR
200	NR
250	NR
250	NR
250	NR
300	NR
300	NR
300	NR

Application demo:

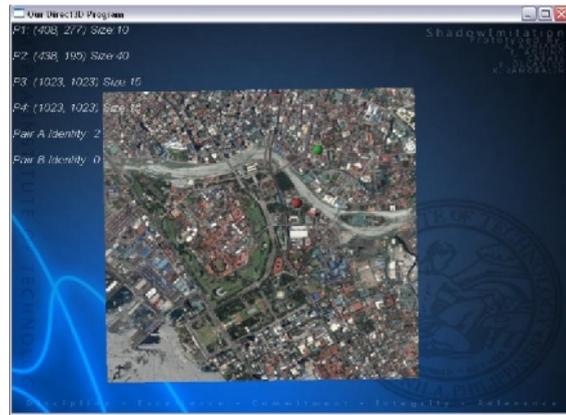
As mentioned earlier, three applications were developed in order to demonstrate the capabilities of the system. Figure 10 A and B shows the Satellite demo application using a map derived from Google Earth program. Figure 10 B shows the zoom-in function, one of the system features, including scroll and zoom-out. Figure 11 A and B shows the Scrapbook demo at the start of the application and after the application respectively. Figure 11 B shows the rearranged pictures utilizing the grab and scroll capabilities. Figure 12 shows the dot pattern using the Infinity symbol. Note that the dotted lines cannot be seen as the gesture is performed. The program will exit if the program interprets the gesture pattern as the Infinity symbol.

Legend:

Red Dot = Left Pinch, Green Dot = Right Pinch

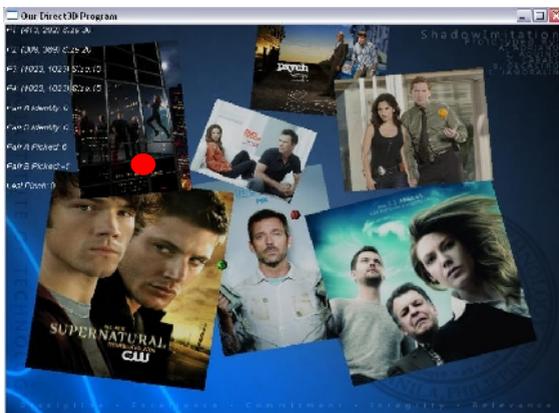


A: Zoom-in



B: Zoom-out

Figure 10 A and B: Gloves Satellite demo application



A: Before



B: After

Figure 11 A and B: Scrapbook demo



Figure 12: Gesture recognition demo

6. Conclusions:

In this paper, a method of interacting with computer system was presented. By using the IS-300Pro Precision Motion Tracker and gloves with infrared LEDs, a prototype was developed that recognizes gestures for basic computer commands like zoom, rotate, scroll. This was demonstrated on the Satellite and Scrapbook demo application. The prototype also recognized gestures for corresponding to commands like opening Internet Explorer or terminating the application by using the Infinity symbol. The prototype was able to recognize gestures corresponding to alphanumeric characters as well as user-defined characters.

Experiments showed that the limitations of optical recognition still exist as far as current technology is concerned. Test results showed that as the distance increases between the infrared light emitters and the sensor, correct interpretation of these gestures decreases. Correspondingly, the complexity of the gestures performed also affects the interpretation to corresponding commands.

References:

- [1]Lementec, J., & Bajcsy, P. (2004), *Recognition of Arm Gestures Using Multiple Orientation Sensors: Gestures Classification*, IEEE Intelligent Transportation Systems Conference.
- [2]Pei Jia, Huosheng H. Hu, Tao Lu, Kui Yuan, *Head gesture recognition for hands-free control of an intelligent wheelchair*, Industrial Robot: An International Journal,2007Volume 34,Issue 1,p 60 – 68, Emerald Group Publishing Limited
- [3]Mr. Pedro Neto, Prof. J. Norberto Pires, Prof. A. Paulo Moreira, *High-level programming and control for industrial robotics: using a hand-held accelerometer-based input device for gesture and posture recognition*, Industrial Robot: An International Journal,2010Volume 37 Issue: 2,Emerald Group Publishing Limited
- [4]Marco Leo, Tiziana D'Orazio, Paolo Spagnolo, Arcangelo Distante, *Automatic video surveillance using statistical analysis of temporal posture sequences*, Sensor,2006 Volume 26 Issue 4,p301 – 311, Emerald Group Publishing Limited

- [5]J.H. Abawajy, *Human-computer interaction in ubiquitous computing environments*, International Journal of Pervasive Computing and Communications, 2009, Volume 5 Issue 1, p61 – 77, Emerald Group Publishing Limited
- [6]Atsushi Shimada, Madoka Kanouchi, Daisaku Arita, Rin-Ichiro Taniguchi, *Robust human posture analysis using incremental learning and recall based on degree of confidence of feature points*, International Journal of Intelligent Computing and Cybernetics, 2009, Volume 2 Issue 2, p304 – 326, Emerald Group Publishing Limited
- [7]Md. Hasanuzzaman, T. Zhang, V. Ampornaramveth, H. Ueno, *Gesture-based human-robot interaction using a knowledge-based software platform*, Industrial, 2006, Volume 33 Issue 1, p 37 – 49, Emerald Group Publishing Limited
- [8]Do Jun-Hyeong, Jung Jin-Woo, Sung hoon Jung, Jang Hyoyoung, Bien Zeungnam, *Advanced soft remote control system using hand gesture*, Mexican International Conference on Artificial Intelligence, 2006
- [9]K. Ouchi, N. Esaka, Y. Tamura, M. Hirahara, M. Doi, *Magic Wand: an intuitive gesture remote control for home appliances*, International Conference on Active Media Technology, 2005 (AMT 2005), 2005
- [10]Thomas Schlomer, Benjamin Poppinga, Niels Henze, Susanne Boll, *Gesture Recognition with a Wii Controller*, Proceedings of the 2nd international Conference on Tangible and Embedded interaction, 2008
- [11]A Erol, G Bebis, M Nicolescu, RD Boyle, X Twombly, *Vision-based hand pose estimation: A review*, Computer Vision and Image Understanding Volume 108, Issues 1-2, October-November 2007, Pages 52-73 Special Issue on Vision for Human-Computer Interaction, doi:10.1016/j.cviu.2006.10.012.
- [12]Stenger B, Thayananthan A, Torr PH, Cipolla R: *Model-based hand tracking using a hierarchical Bayesian filter*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 28(9):1372-84, Sep 2006.
- [13]Matthias Rehm, Nikolaus Bee, Elisabeth André, *Wave Like an Egyptian - Accelerometer Based Gesture Recognition for Culture Specific Interactions*, British Computer Society, 2007
- [14]Alejandro Jaimesa and Nicu Sebe, *Multimodal human-computer interaction: A survey*, Computer Vision and Image Understanding Volume 108, Issues 1-2, October-

