# 3D Spatial Sensor

By

**Tan Shern Shiou, 6517**

Final Report submitted in partial fulfillment of
the requirements for the
Final Year Project

Bachelor of Technology (Hons)
(Information Communication Technology)
JANUARY 2008

Universiti Teknologi PETRONAS
Bandar Seri Iskandar
31750 Tronoh
Perak Darul Ridzuan

# Final Year Project

Student: Tan Shern Shiou

Student ID: 6517

Supervisor: Low Tan Jung

CERTIFICATION OF APPROVAL

3D Spatial Sensor

By

**Tan Shern Shiou**

A project dissertation submitted to the
Computer and Information Science Programme
Universiti Teknologi PETRONAS
in partial fulfilment of the requirements for the
Bachelor of Technology (Hons)
(Information and Communication Technology)

Approved by,

(Mr. Low Tan Jung)

UNIVERSITI TEKNOLOGI PETRONAS
TRONOH, PERAK
January 2006

CERTIFICATION OF ORIGINALITY

# 3D Spatial Sensor

This is to certify that I am responsible for the work submitted in this project, that the original work is my own except as specified in the references and acknowledgements, and that the original work contained herein have not been undertaken or done by unspecified sources or persons.

Tan Shern Shiou

**Acknowledgement**

3D Spatial Sensor is one of the ideas that arose in the 8 months at MIMOS Berhad. During the internship period, I worked with few other colleagues who teaches me, exposed me with knowledge especially in research techniques. But my uttermost gratitude is given to Dr Hon Hock Woon which is my supervisor during my internship. With his assistance, we did not only come out with the idea of 3D Spatial Sensor but he also helped in patenting this idea in MIMOS. Besides, his help and support to me that extend till these days are greatly appreciated. His knowledge and experience are exceptionally inspired and enrich my growth as a student and a researcher.

During my final year where I conduct my research work on this final year project in Universiti Teknologi Petronas, I have worked with Mr Low Tan Jung who is my supervisor for this particular project and contributed in assorted ways in this project deserved a special mention. I would like to express my gratitude to Mr Low especially being patience and helpful during the whole process. I am grateful in every possible ways and hopefully there will be more of our collaboration in the future.

To all of my friends who helped in this project especially Lee Kee Foo, Beh Sea Min, Ong Wei Shen, Tan Wei Zhu, Mahendran Balakrishnan and Aizat Kamaruzaman. Your supports and comments helped me sail through exhibitions and presentation with ease. To others that are not mentioned, thanks for your patience. If there are misunderstanding and miscommunication, I would like to apologize here. I appreciated your supports and hopefully our friendship will be forever.

I would like to record my gratitude to my family members who always support me and educate at their best since I was a toddler. Where am I without my family members? Thanks for paving my life and giving me the best learning environment that make me a better person now.

## Abstract

This project entitled 3D spatial sensor which is a sensor that is designed for sensing touch coordinates in 3 dimensions for 3D application such as 3D modeling and human computer interface. The main idea of this project is to develop 3D sensor that can be used by everybody with minimal learnability and cost.

3D sensor is important to this world because our real world is a 3 dimensions world hence our interaction with computer has to be in 3 dimensions so that the real data can be inputted into the computer for further processing. However, many of the sensors for human computer interaction is limited to 2 dimensions such as keyboard, mouse and touch screen device, hence the real interaction between human and computer is limited by those devices.

In this project, the development of 3D spatial sensors will be based on a novel method which make used of 2 cameras for 2 different spatial planes and a computer to calculate and derived the 3D coordinates from both the image frames from the 2 cameras. From a single touch 3D spatial sensor, additional efforts and resources will be used to develop a multiple touch 3D spatial sensors.

Lastly, multiple touch 3D spatial sensors will be incorporate into different applications to demonstrate the abilities of these sensors. Hopefully the development of this device can benefit human and enhance the experience of human computer interaction.

# Table of Contents

# List of figures , diagram and images

**Diagrams**



**Figures**



**Images**

## Abbreviations

| | | |
|---|---|---|
| HMM | - | Hidden Markhov Model |
| IPLimage | - | Intel Image Processing Library Image |
| OpenCV | - | Open Computer Vision |
| 2D | - | 2 Dimensions |
| 3D | - | 3 Dimensions |
| FTIR | - | Frustrated Total Internal Reflection |
| API | - | Application Programming Interface |
| CCD | - | Charge Coupled Device Image Sensor |
| CMOS | - | Complementary Metal Oxide Semiconductor Image Sensor |
| Fps | - | Frame per second |
| USB | - | Universal Serial Bus |
| CPU | - | Central Processing Unit |

# 1. Introductions

## 1.1 Title

3D Spatial Sensor

## 1.2 Problem Statement

2D touch sensors are commonly used nowadays for user interface with electronic devices because of its intuitive interfaces and high learnability features. Due to its 2 dimensions coordinates, there is a lot limitation to enhance the real user interface because our world is in 3D. With 2D touch, we need a touch surface for inputting the touch location. Besides, 2D touch does not allow us to input the depth of each touch and does not cover Z plane.

To overcome these problems, 3D touch sensor can be produce to let user input its 3D coordinates. But current 3D sensors have few limitations.

## 1.3 Objective / Scope of Studies

1. Develop a hardware prototype for image capture
2. Derive a project flow using prototyping method to build the prove of concept
3. Suggest potential application that is useful for this idea
4. Research on few image processing method to find the most suitable for this prototype
5. Develop a simple sample application to demonstrate the ability of 3D Spatial Sensor

## 1.4 Short Summary of Project

This novel project idea is already disclosed under my name and my internship supervisor during my industrial internship period parked on MIMOS Berhad. Using 2 cameras where a camera is positioned at Z,Y plane capturing image data for (Z,Y) coordinate while another camera will be positioned at X,Z plane capturing image data for (X,Z) coordinate.

Both the image data will be going through background and foreground segmentation using image processing function to retrieve the touch blob for both Z,Y plane and X,Z plane. The (Z,Y) coordinate can be calculated from Z,Y plane by calculating the distance with the reference map. The same goes to the X,Z plane too. To get the 3 dimensions coordinate, we can match the Z,Y coordinate and X,Z coordinate to get X,Y,Z coordinate through the matching processing of the common Z point of both (Z,Y) and (X,Z).

An application for this 3D spatial sensor can be developed to demonstrate the feasibilities and the advantages of 3D spatial sensor over 2D touch sensor or any input user interfaces such as keyboard or keypad. A application is proposed is to develop a user interface for security application such as ATM where 3D spatial sensor act as a user interface for replacing the keypad of ATM. This application can avoid the users to directly touching the keypad which can be noticed by outsiders with bad intention and leaving fingerprint on the keypad that can be recognize for PIN code.

3D spatial sensor also can be develop for applications for graphical or video industry because 3D spatial sensor does not need motion sensor for them to get the 3D coordinate for them to capture 3D data. As for the result, this idea can help the 3D designer to input 3D coordinates easily.

For the optional enhancement of this project, idea for multi-touch spatial sensor can be developed by training the Hidden Markov Model (HMM) to recognize each of the object blobs and match the object blobs for both of the planes which output every (X,Y,Z) coordinates for every blobs detected.

## 1.5 Limitation of Current 3D sensor

Usage of Markers – Marker is not user friendly and it is not practical for minute touch such as finger touch. Besides, marker does not complaint with ubiquitous computing where user can use it with minimal learnability.

Physical device – Some 3D sensor utilise a physical device like a ball or a pen for user to input 3D coordinates. This device requires the user to familiarise and learn to hold it properly.

Single touch – Most 3D sensor currently only can output a single 3D coordinates which is a big limitation.

## 1.6 Solution

3D spatial sensor that utilise machine vision to predict every peak of object and matches both of the planes for a full multi touch 3D coordinates. The main advantage of using machine vision is the user do not have to wear any mocap or any device to track the position of the touch hence it increase the learnability and reduce the cost as the development of the whole sensor is only at the machine itself. Besides, the advancement of machine vision and image processing in current trends with the help of powerful computer and camera, the whole system is more efficient and easier to develop.

# 2 Prior Art

## 2.1 Existing solution

Phantom Haptic Device

A pen like shaped device connected to a base that mainly used for 3D modelling where the location of the tip of the pen will be the reference point for 3D coordinates as the user move the pen.

Spaceball

A ball that is place like a trackball for user interface that capture coordinates and predict the space position mainly for 3 dimensions Graphical User Interface such as 3D deskstop.

Ultrasonic sensor

Ultrasonic sensors are placed to sense touches that is above the sensor hence creating 3D coordinates based on reflected length of ultrasonic and placement of detected ultrasonic sensor.

Mocap with Marker

Motion capture sensor that utilise the ray sensitive markers that is placed on body for capturing 3D coordinates for motion.

## 2.2 Less related work

FTIR sensor

A multitouch 2D sensor that uses Infra-Red ray refraction on a glass to produce touch spots for 2D multitouch sensors.

## 2.3 Summary of Literature Review

Cost

3D Spatial Sensor will be minimal cost because it needs only 2 webcam and a CPU to perform the function of outputting 3D coordinates. For Mocap, the device is expensive because of the marker and the usage of several large and expensive cameras to capture and derive the 3D coordinate.

Marker / Equipment

3D Spatial Sensor is unique because of it do not need the usage of the marker to produce 3D coordinate hence it is closer to human, easy to employ and high learnability. Without the usage of marker, easily every human can use which make it suitable to be 3D interface for human. With marker or other device, the people need to learn and adjust so that they can use the interface. Other than that it is harder to setup and expert is required to get a good output.

Setup Easiness

3D Spatial Sensor do not require expertise in setting up the system and it only need a small confinement space with specific environment requirement. While compare to Mocap system, the software itself is very complicated and setting up the hardware need a large space and assistance from expert.

Accuracy

Comparing 3D Spatial Sensor and other existing, this innovative method of producing 3D coordinate is still new. Hence, the level of accuracy is not very high due simple image processing part and less optimum CPU usage. But, if this technology is better develop, the accuracy will be comparable to the existing 3D user interface.

# 3. Methodology

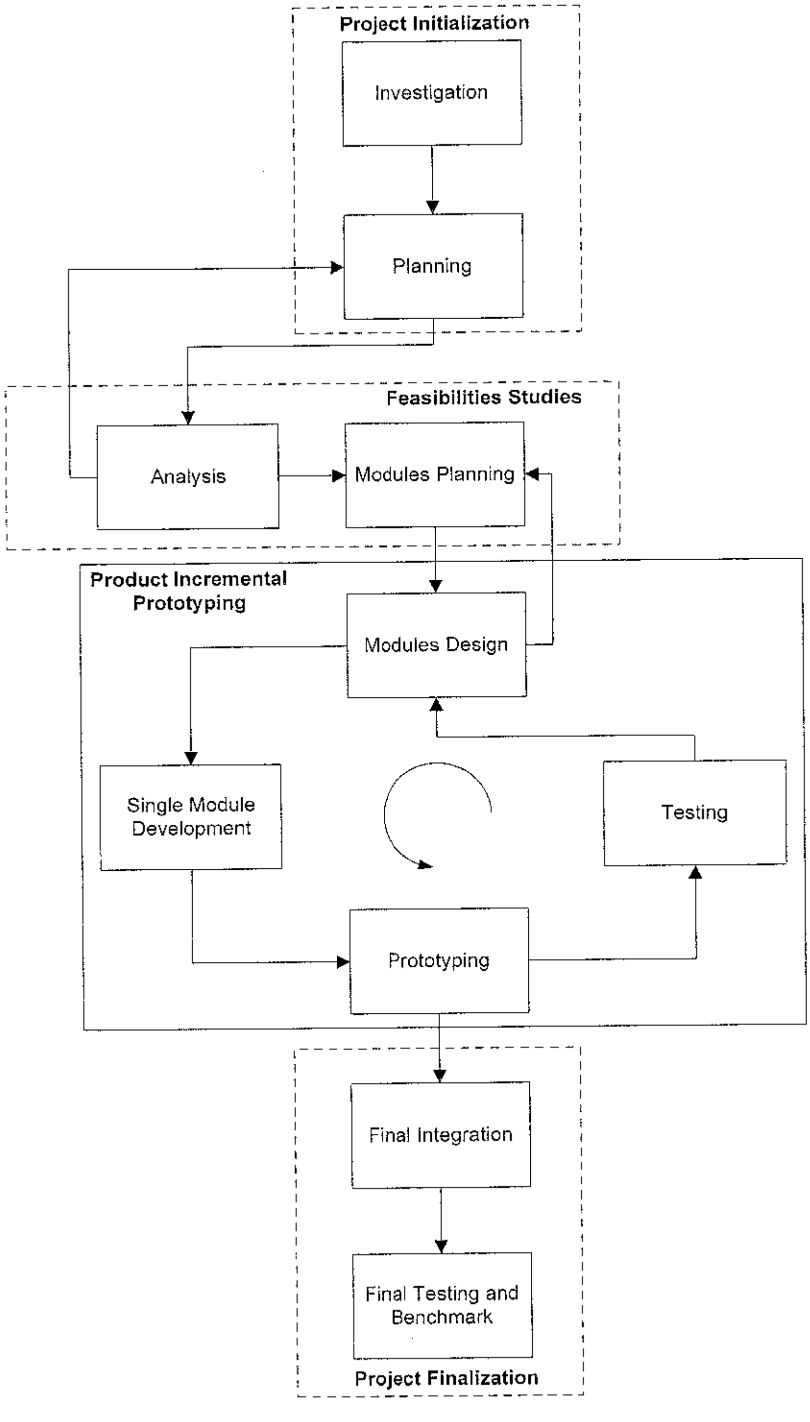



Diagram: Methodology

Methodology that will be used in this project is a mixture of modules development, prototyping and spiral development model. These methodologies are chose because all of them have their own benefit.

- Modules Development Model – Able to develop functional module one by one and test module by module to make sure the each module is workable
- Prototyping – Able to visualise quick prototype before the actual product is being produced and go back if the prototype do not produce expected result
- Spiral Development – Able to go incremental development and plan for the next level of development

The project starts with project initialization where it involves investigation and planning. Investigation will be problem analysis, prior art investigation and patent searching. Planning includes the task to plan to conquer the whole project until finalizing the project.

The next steps will be feasibilities studies where analysis and modules planning are taken into action. Analysis involves the feasibilities studies on the method proposed and searching the most feasible ways to fully utilise the method and produce the best result. Modules planning is followed by the analysis where planning for every module that will be develop for the whole project.

Product incremental prototyping involve module design, single module development, prototyping and testing. Each and every module will go through these 4 steps in order to pass the stage. Module design is focusing on the overall structure and blueprint for a single module that will be develop in the later stage. After development, it will be prototype with the previous module for testing to pass a certain expected result.

Final integration is to integrate and packed all the modules develop in the product development stage. Final testing and benchmark will be test the overall status and benchmark it for analysis.

# 4. Theory and Technique

This unique technique which is 3D spatial sensor that make use of image processing on 2 camera image and produce a predictive 3D coordinate through matching of both planes.

## 4.1 Technique

The main concept behind this 3D spatial sensor is the use of 2 cameras to capture both spatial planes XZ and ZY planes. After capturing the plane, the computer will segment out the background and foreground in order to get the intended foreground for further analysis. The foreground of the images will be turned in blob after image processing and enhancement for spatial sensing purposes. The blob will be calculated and determine the position of the touches for both planes and then mathematically calculate to combine the coordinates from both planes. The output of the calculation will be a 3 dimensions coordinate (X,Y,Z).

## 4.2 Approach

### 4.2.1 Single touch spatial sensor

1. Capture images from 2 different cameras connected.
2. Background and foreground segmentation on both the camera images to get the foreground which is the fingers.
3. Get the peak point of each foreground in coordinates.
4. Match both coordinates by using reference common point from common plane.

### 4.2.2 Double touch spatial sensor

1. Capture images from 2 different cameras connected.
2. Background and foreground segmentation on both the camera images to get the foreground which is the fingers.

3. With mathematical function, 2 peaks of the blob to determine 2 touches.
4. Get the peak point of each foreground in coordinates.
5. Match both coordinates by using reference common point from common plane.

### 4.2.3 Multi-touch spatial sensor

1. Capture images from 2 different cameras connected.
2. Background and foreground segmentation on both the camera images to get the foreground which is the fingers.
3. Using skin feature extraction to get the skin part of the foreground blob.
4. Extract the blob by using contour extraction.
5. With mathematical function, peak and valley of the finger is identified.
6. Intelligently match both peaks from both planes to get 3D coordinates.

# 5. Equipment / Tools

Procurement of tools and equipment for 3D Spatial Sensors:

1. Visual Studio .NET 2005

   To develop the software for machine vision and image processing while controlling the camera frames and deduces the coordinate using C language.

2. OpenCV

   Open source image processing library which include background and foreground segmentation, mathematical function for blob analysis and much more.

3. 2 webcams

   Webcams is a cheap alternative beside camcorder to capture the image of the touch on the spatial planes. Each webcam represent different spatial plane.

4. Frame

   To hold the webcams at the correct position after the calculation to determine the optimum space for spatial sensing.

# 6. Discussion

## 6.1 Project Initialization

Investigation on problem statement has been done where the problems of 3D touch sensors in the market are identified. To build a good human computer interface and improve the touch sensing device, the existing problem must be tackle. A new and novel idea generated using image processing techniques and cameras to overcome the existing problem giving us a 3D touch sensor without needed for human to wear marker with high learnability and usability. Besides, this idea of improved touch sensing interface is aimed to be highly integrable with other machine.

After the problem statement and research has determined in the investigation phase, planning is the stage where we layout the future course and come out with a schedule which include step by step progress and milestone in order to achieve the result. A methodology is introduced to accommodate this special modular development.

## 6.2 Feasibilities Studies

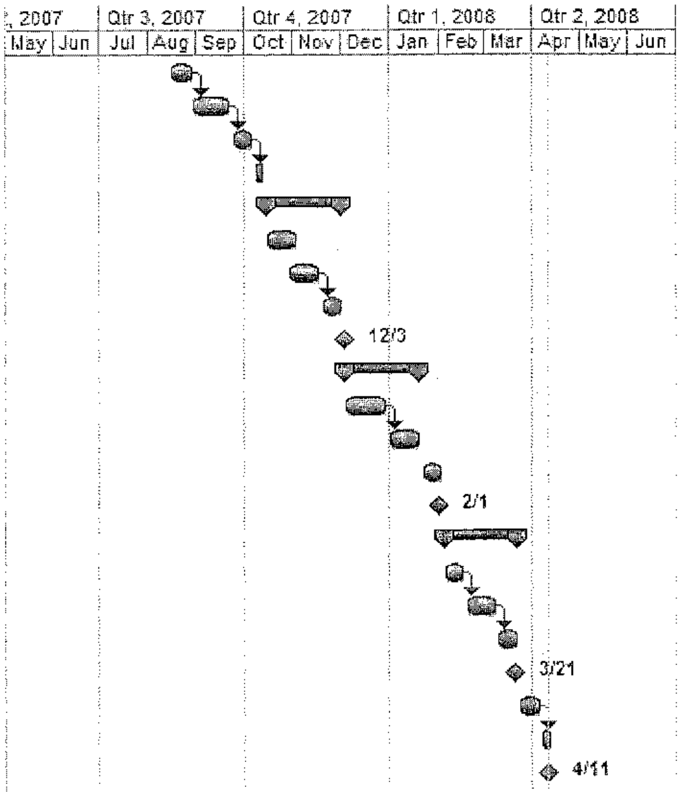
Feasibilities studies stage includes analysis and modules planning. From the analysis phase, the outcome of the analysis is the layout of the schedule and the modules to be worked on for the project. There are two milestone need to be completed before end of this year which is single touch and multiple touch. Conceptual diagrams have been illustrated to visualize the idea in figure format.

After modules planning, a double touch milestone is inserted between single touch milestone and multiple touch milestones because it will ease the transition from single touch to multiple touch and double touch sensor is easier to develop yet fulfil the requirement of multiple touch module. Modules planning will be updated and revisit every time each modules has finished so that the project has some flexibility in terms of deliverables.

## 6.3 Schedule

The whole schedule has been planned for two semesters from the start of the project until the end.

| Task Name | Duration | Start | Finish |
|---|---|---|---|
| Investigation | 10 days | Wed 8/15/07 | Tue 8/28/07 |
| Planning | 18 days | Wed 8/29/07 | Fri 9/21/07 |
| Analysis | 10 days | Mon 9/24/07 | Fri 10/5/07 |
| Modules Planning | 5 days | Mon 10/8/07 | Fri 10/12/07 |
| **Image Capture Module** | **35 days?** | **Mon 10/15/07** | **Fri 11/30/07** |
| Single USB Webcam | 15 days? | Mon 10/15/07 | Fri 11/2/07 |
| Double USB Webcam | 15 days? | Mon 10/29/07 | Fri 11/16/07 |
| Module Testing | 10 days? | Mon 11/19/07 | Fri 11/30/07 |
| Image Capture Milestone | 0 days | Mon 12/3/07 | Mon 12/3/07 |
| **Image Processing Mod** | **35 days?** | **Mon 12/3/07** | **Fri 1/18/08** |
| BG FG Segmentation | 20 days? | Mon 12/3/07 | Fri 12/28/07 |
| Coordinate Module | 15 days? | Mon 12/31/07 | Fri 1/18/08 |
| Module Testing | 10 days? | Mon 1/21/08 | Fri 2/1/08 |
| Image Processing Mileston | 0 days | Fri 2/1/08 | Fri 2/1/08 |
| **Image Display Module** | **35 days?** | **Mon 2/4/08** | **Fri 3/21/08** |
| Coordinate Input | 10 days? | Mon 2/4/08 | Fri 2/15/08 |
| OpenGL Programming | 15 days? | Mon 2/18/08 | Fri 3/7/08 |
| Module Testing | 10 days? | Mon 3/10/08 | Fri 3/21/08 |
| Image Display Milestone | 0 days | Fri 3/21/08 | Fri 3/21/08 |
| Testing and Benchmarking | 10 days? | Mon 3/24/08 | Fri 4/4/08 |
| Finalization | 5 days? | Mon 4/7/08 | Fri 4/11/08 |
| Project Milestone | 0 days | Fri 4/11/08 | Fri 4/11/08 |



### 6.3.1 Schedule Items

- Investigation
- Planning
- Analysis
- Modules Planning
- Device Frame Module
- Image Capture Module
    - Single USB Webcam
    - Double USB Webcam
    - Module Testing
    - Image Capture Milestone
- Image Processing Module
    - Background and Foreground Segmentation
    - Coordinate Module
    - Module Testing
    - Image Processing Milestone
- Image Display Module
    - Coordinate Input
    - OpenGL Programming
    - Module Testing
    - Image Display Milestone
- Testing and Benchmarking
- Finalization
- Project Milestone

## 6.4 Device Frame Module

The purpose of the device frame is to hold both of the cameras in right angle because the camera must put in a right angle to capture accurate coordinate value from XY plane and XZ plane. Device frame has built according to the design of Diagram: Device Frame with 2 bar each measured 40cm and being held with a based strong enough to hold the weight of the cameras and both the bars.

Progress:

The device frame is built with two plastic casing as both the bars with cover that can slide to hold both the webcams. The purpose of the slide case that hold the webcams is to make sure both the webcams are adjustable in terms of height so that the optimum field of view of each cameras are able to be determined.

The device frame is measured 40cm each bars because 40cm is a very good height for the webcam with narrow field of view to capture for this sensing project. As the field of view of camera is pyramid in shape, the further away from the camera, the effect of pyramid slope is not significant hence the overlapping region of both the cameras will be slightly like a cube. Cube shape of field of view is essential because the normal Cartesian 3D coordinate is meant for cube. The illustration of detected area is illustrated in Diagram: Detected Area.

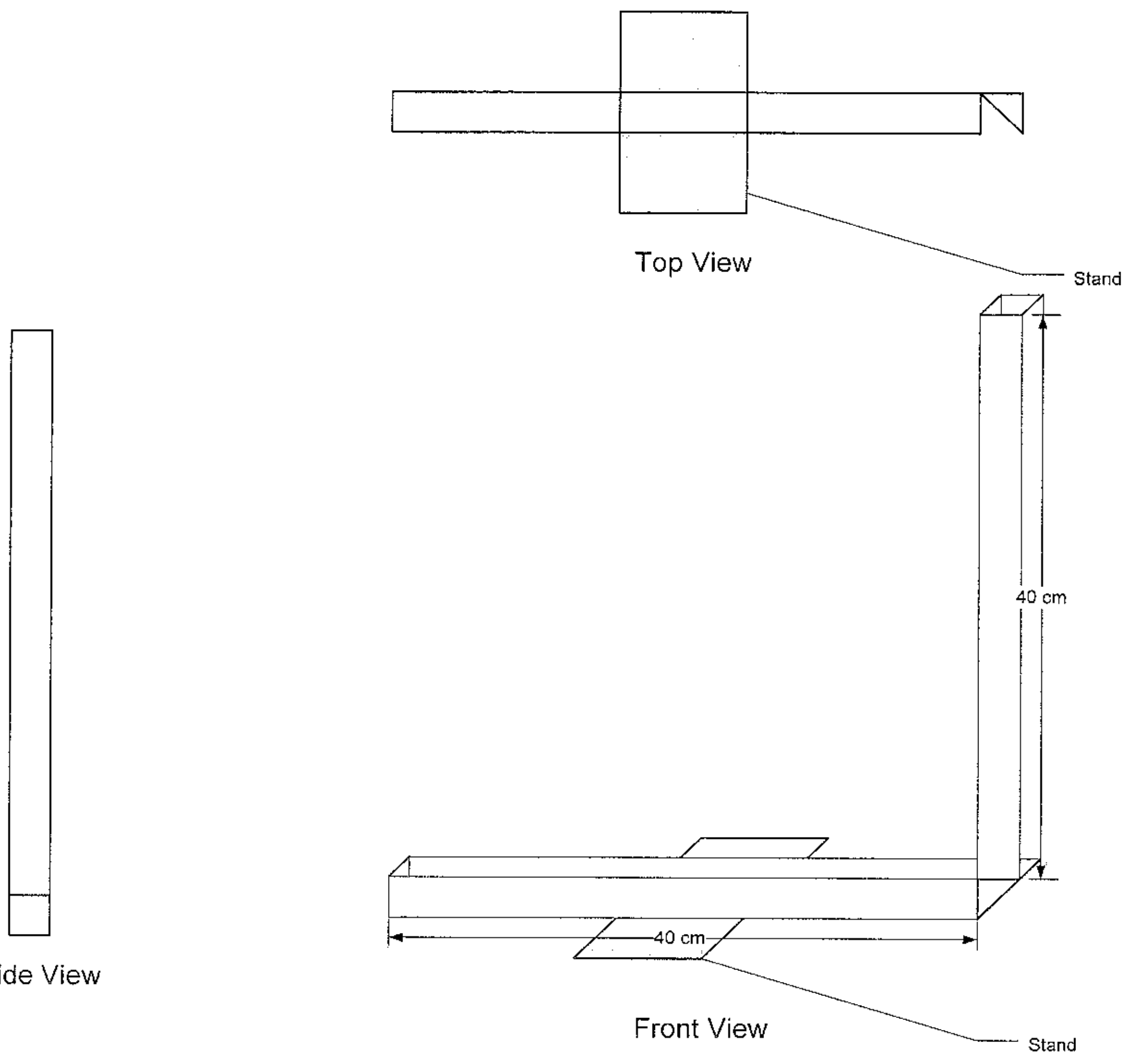



Diagram:
Device Frame

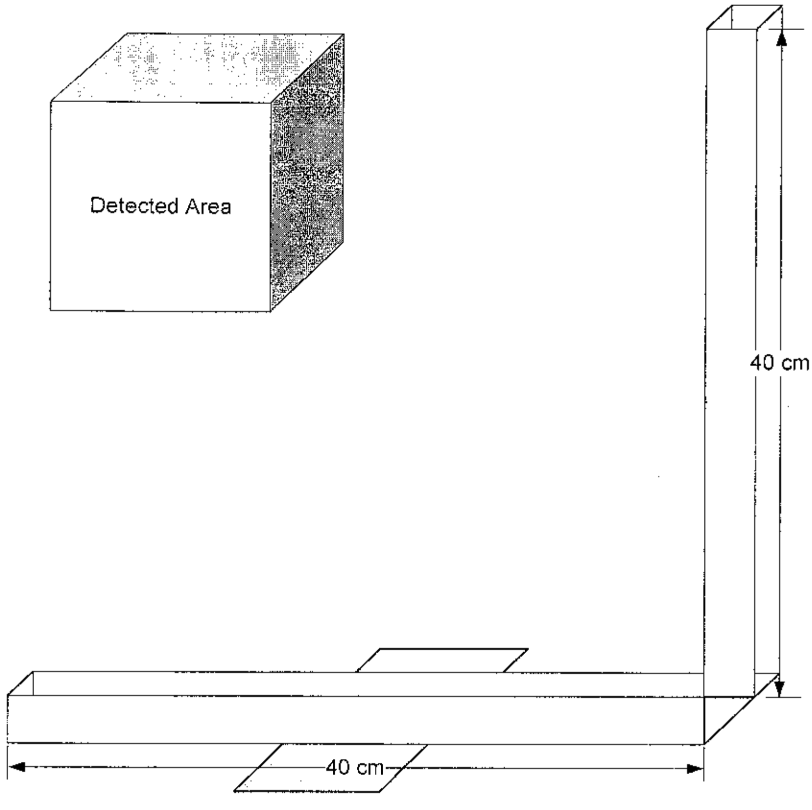



Diagram:
Detected Area

The diagram above illustrate about the detected area from the overlapping field of view of both the webcam position in right angle. The detected area is relatively small with about $5cm^3$ and floating in the air. To operate this 3D Spatial Sensor, we must place our finger inside the detected area.

## 6.5 Image Capture Module

Using Video for Windows (VFW) to grab and retrieve the frames from both the cameras through USB ports. The frame grabbing must able to match the maximum frame rate where the web cam and the bandwidth of USB port can offer to get the highest quality of touch sensing possible for further level. This image capture module is very important because the further level of this sensing application involve image processing where it will reduce the frame rate when it reaches the user.

Goals:

1. Capture 2 frames concurrently from 2 webcams through USB ports
2. To achieve 30 fps (maximum web cams frame rate)
3. Able to convert the raw image data into IPLimage for OpenCV processing

Progress:

This module is nearing the ending phase of a module where this module is functioning correctly with concurrently capturing 2 frames at one time from 2 webcams through USB ports. Both the raw image data from webcam 1 and webcam 2 are able to convert into IPLimages so that image processing can be done with OpenCV library.

But the frame rate in this level is not satisfactory where the object movement in both the frames are not smooth and do not get pass the 20fps mark. Further tweaking is needed to enhance and improve the frame rate including changing the capturing concept into multiple call-backs per second.

After changing to multiple call-back per second, the result of image capturing is improved but it still do not achieve the goal of this image capture module which is

## 6.6 Image Processing Module

Image processing module consists of 2 main parts which is Background and Foreground Segmentation and Coordinate Module.

### 6.6.1 Background and Foreground Segmentation

Using OpenCV image processing library, mathematical function and calculation are used with the frames captured to determine the background and foreground of the frames. The mathematical function will compare the background and the movement changes in the following frames where several requirements need to be met before the background and foreground segmentation module can function correctly.

Requirement of module:

- Static web cams position
- Determine the background before segmentation started
- No background object move after the determination

Goals

1. Stop point to determine the background before segmentation begin
2. Segment out the foreground leaving the background in black and white format
3. Determine the best condition for segmentation progress

Progress

From the IPLimage converted from the first module, background and foreground segmentation can be done to a single frame. Further implementation will be done to both the frame when the first segmentation is improved. A stop point to get the first frame as the background is implemented where the user need to press on a key indicating to the computer the background has been determined and the segmentation can start.

Work has been done to determine the colour combination to get the best quality of segmentation with reduced noise and best frame rate. Y'GbGr colour mode can be used for the frames before the background and foreground segmentation because this colour mode proved reduce the noise compare the normal RGB colour mode while still maintaining the foreground segmented.

### 6.6.2 Coordinate Module

Coordinate module will create 2 coordinate layout for both foreground segmented to differentiate the XZ plane and ZY plane. Due to the difference between pixel coordination and Cartesian coordinate, coordinate module is essential so that we can determine the location of each touch and translate into normal Cartesian 3D coordinate.

To determine the point for the touch, in coordinate module must have the peak detector to detect the peak of the touch so that the peak point will be the coordinate of the touch.

Goals:

1. Peak detector
2. XZ plane and ZY plane coordinate layout

Progress:

The coordinate module is completed with the peak detector for each plane. With peak detector, the computer computes the peak of each blob (finger) and provides the 2 dimensional coordinate for the plane. The peak detector also circle the peak to provide an indicator of the peak detected so that the developer can determine whether there is error exist in the detected peak.

As for the reference for other plane, a line is drawn to indicate the reference axis is drawn to show the matching coordinate. Both the plane must match to generate the 3D coordinate. After some calibration between the cameras, the matching point can be found.

## 6.7 Image Processing Milestone

Requirements:

- Output 3D coordinate frame by frame
- Display foreground, background.
- Label both foreground and include a point detected position.

Image Processing Milestone is reached with the entire basic requirement is done. After the testing phase of this first prototype, then this project will progress to the next milestone which is Image Display Milestone where the whole prototype will be completed.

## 6.8 Image Display Module

The image process part of the device can be toggle to let the developer see the internal process of the whole device to debug and check the process flow. The image display module is the mainly for the developer to visualize the image processing.

Image display module is essential because through this image display module, the developer able to test and integrate other function to make full use of this sensor. Besides, computer can calculate the performance rate of this device so that the developer can make some tweaking. The function of 3D environment in this image display module is to display the area of the finger touch to determine whether the sensing area is correct through rough eye inspection of finger location and the pointer location in the 3D environment displayed on the computer monitor.

Requirement of module:

- 2 window for both planes with indication of detected area
- Displaying match coordinate accordingly at a console window
- Determine the error rate and frame rate for performance checking.
- Create a 3D environment through OpenGL / DirectX

### 6.8.1 OpenGL Programming - Keypad Emulation Module

Keypad emulation is an application for the user to demonstrate the ability of this 3D sensor. The purpose of this module is to let the user touch on air as if they are touching the keypad and generate number based on where they are touching.

Requirement of module:

- Determine and display the position of finger
- Design and specify the place of button / icon
- Perform simple function for the user to interact

Goals

1. Draw the coordinate and button in 3D
2. Approximately sensing the pressed button through calibration
3. Translate the touch area into monitor screen

## 6.9 Image Display Milestone

Visualization part is always the last part of an image processing application where the monitor or any display device will display useful information to interact with the user hence it complete the whole package as an interactive application and serve the purpose of this whole project.

# 7. Failure

After testing the first prototype, a lot unexpected failure can be found. Therefore, the project scope has to be reduced so that the main objective can be reach within the timeframe.

*Failure to be corrected*

1. Slow detected frame rate
2. Noise
3. Error of peak detection
4. Incorrect coordinate matching

## 7.1 Suggestion / Correction

### 7.1.1 Active contour / Snake

Active contour is a method of analysing a moving image where it take the outline from edge detection and determine the contour of the image. From the contour of the image, we can find the intended blob / object which is the finger position in this case.

Benefit

1. Bypass background and foreground segmentation (Improve detected frame rate)
2. Capture the intended blob (Reduce noise)
3. Smoother detected blob

### 7.1.2 Distance Transform

Distance Transform of the pixel in an image is a proven method in gesture recognition and finger detection in most of the image processing application. Pixel blob can be calculated and a pixel heat map can be generated with grayscale value to determine the area where there are more pixels. By using this method, we can get the correct peak more often by determine the right grayscale value and the approximate position of the finger.

### 7.1.3 Enclosure

Enclosure of the whole device is suggested to be made to help in background and foreground segmentation so that the image processing module will perform a better result. The function of the enclosure is to provide a uniform and contrast colour background to ease the background and foreground segmentation calculation process by the computer.

In this project, the enclosure is chosen and made by two pink coloured cardboard which hold up together with tapes to form background for each camera. As a result a crisper image can be formed and the result improved. The accuracy of the detection has risen but it is still at minimum level due to the fact that the noises from the webcams are unavoidable. The enclosure also made the whole detection environment darker forcing the webcam to turn into low light capture condition hence it risen the noise level.

# 8. Future Application

There is a lot application that can be associated using this 3D spatial sensor technology:

### 8.1 Keyboard / keypad Simulation

An on space keyboard / keypad can be developed by setting the key position on a certain square coordinates spaces and when the touch reach the coordinate, a key will be issued out based on the space. This technique is very feasible for high security requirement machine because this type of touching do not leave any fingerprint therefore it does not leave any mark / track on the keypad.

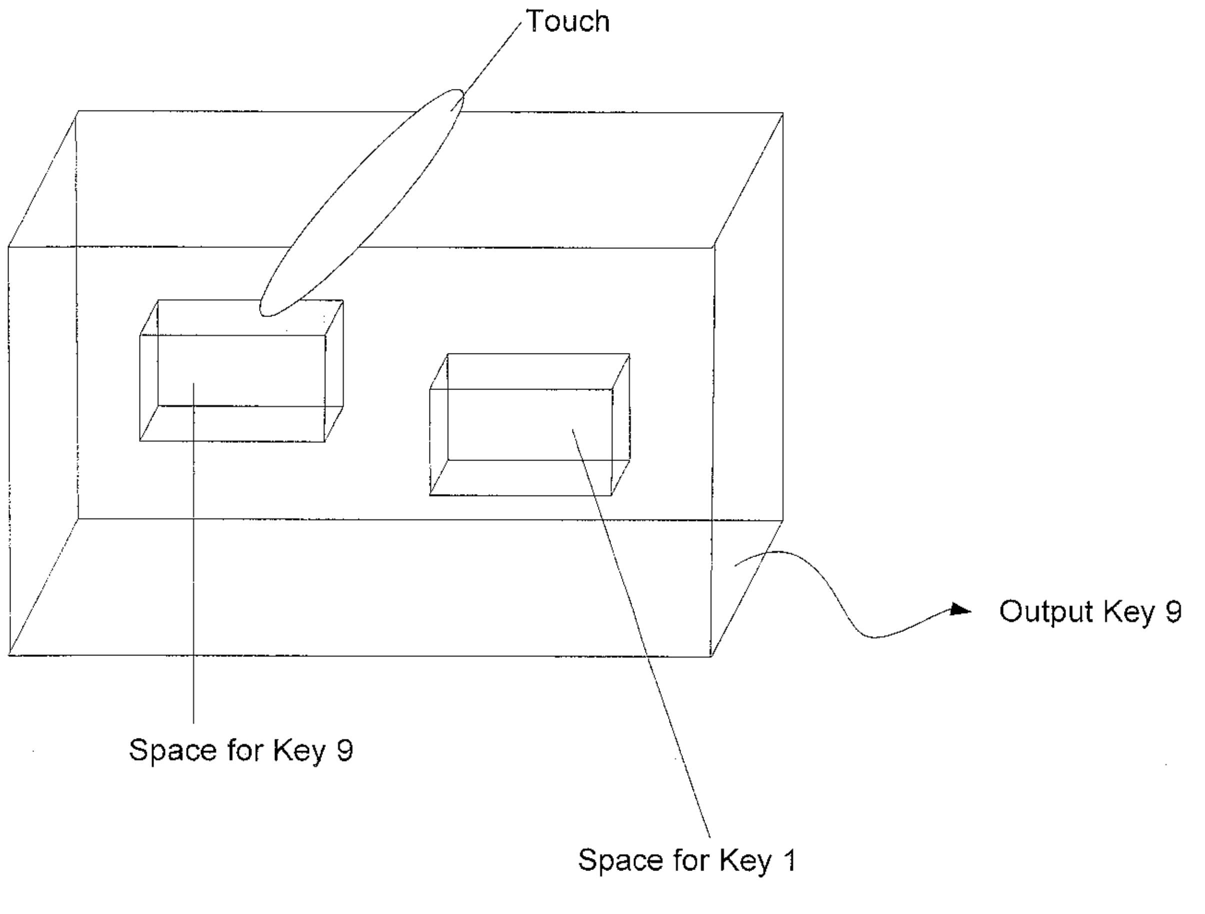



Figure 1: Keyboard / Keypad Emulation Diagram

## 8.2 3D Graphic Modelling

In traditional ways, the coordinates need to be inputted into computer graphic modelling software for rendering by inputting from keyboard or vertices from mouse. This technology can input the vertices by actually touching the spaces on the air because this device actually output the actual 3D coordinates / vertices. Besides, 3D spatial sensor can control the 3D object on screen either pan, tilt or zoom by touching some movement inside the spatial space while viewing on the monitor. Figure 2 illustrate the ability of tilting the 3D object on computer monitor by touching in the spaces in 3D spatial sensor space and the output coordinate will act as reference point of the tilting point of the cylinder.

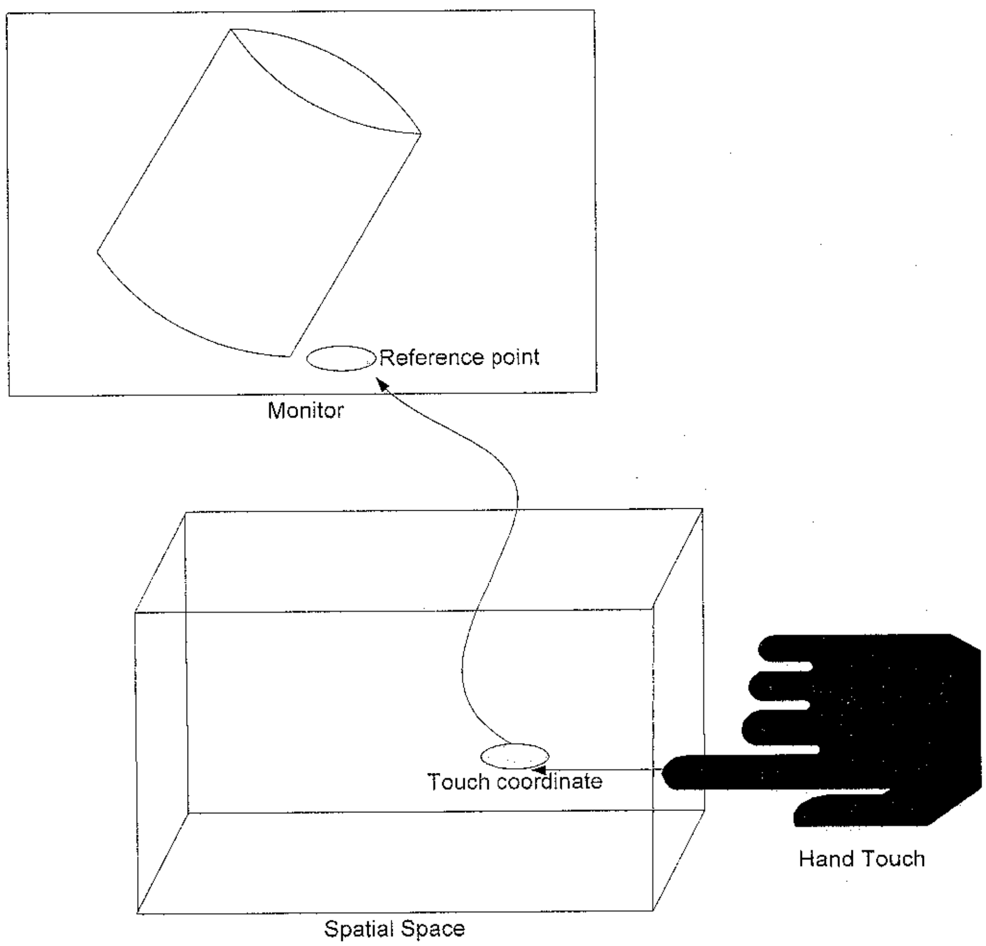



Figure 2: Illustration of Tilting 3D object using 3D Spatial Sensors

## 8.3 3D Gesture Recognition

In the present gesture recognition, it is only made for 2D application where the camera is place on the end and captures the gesture on X, Y plane. This method neglects the Z plane where the gesture does not incorporate the depth of the gesture. By using the spatial sensor, we can record the gesture and because the sensor output 3D coordinates, we will have X, Y, Z planes of gesture where it can enhance the present gesture recognition furthermore.

The gesture recognition can be trained and recorded in the form of neural networks so that the gesture can be detected in other application that accepts the interpretation of the specific neural network. To train the neural network, the inputs and the weight adjustment of the coordinate must be put inside the neural network. The movement of the finger and the resulted 3D coordinates can be the main source of training the neural network.

# 9. Result

The result of this 3D spatial sensor prototype is represented in several windows with 2 windows for image captured from the webcams, 2 windows for segmented image from the image frames captured, a text that states the generated output of the coordinates and a window dedicated for showing the coordinates in boxes.

## 9.1 Image Capture Module Result

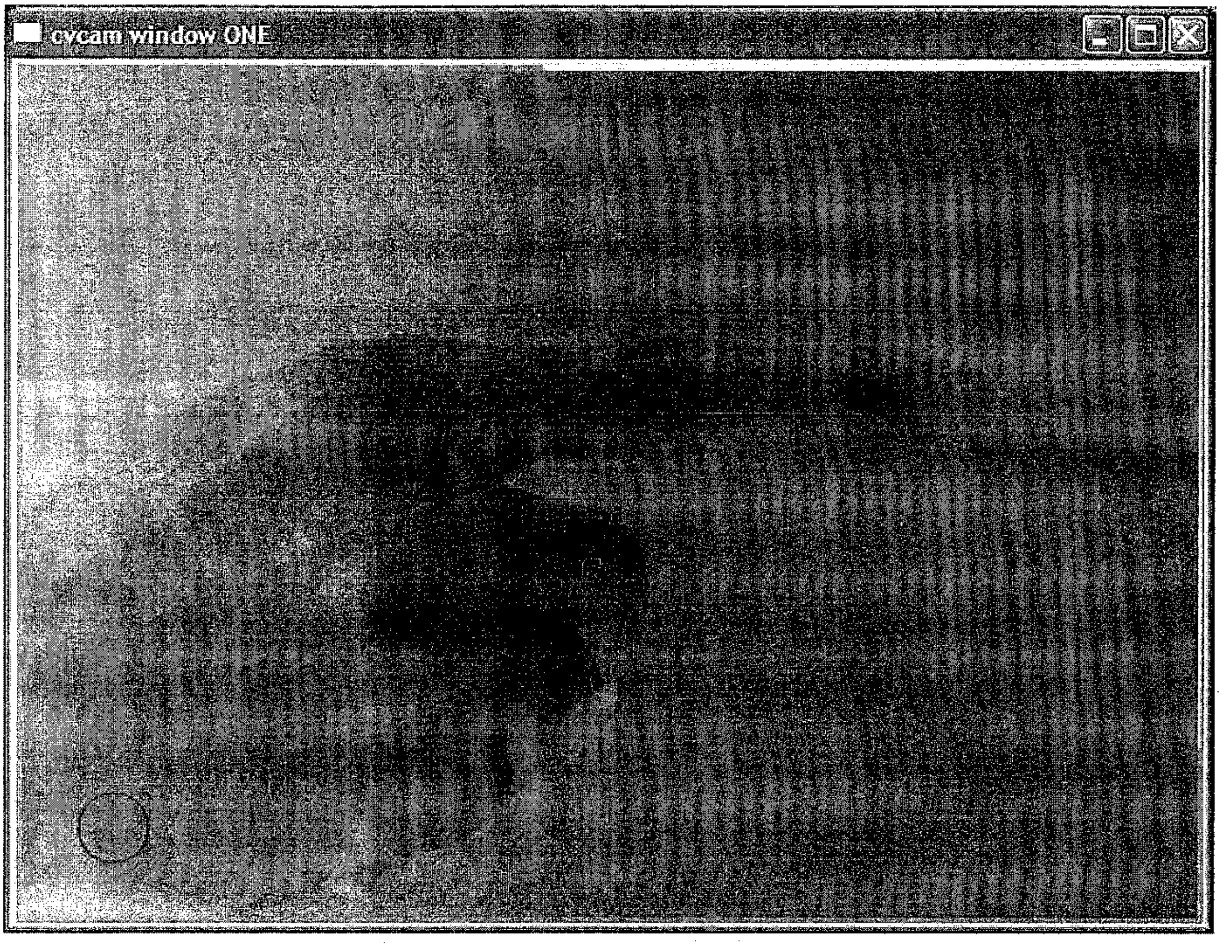



Image 1: Image captured from YZ Plane

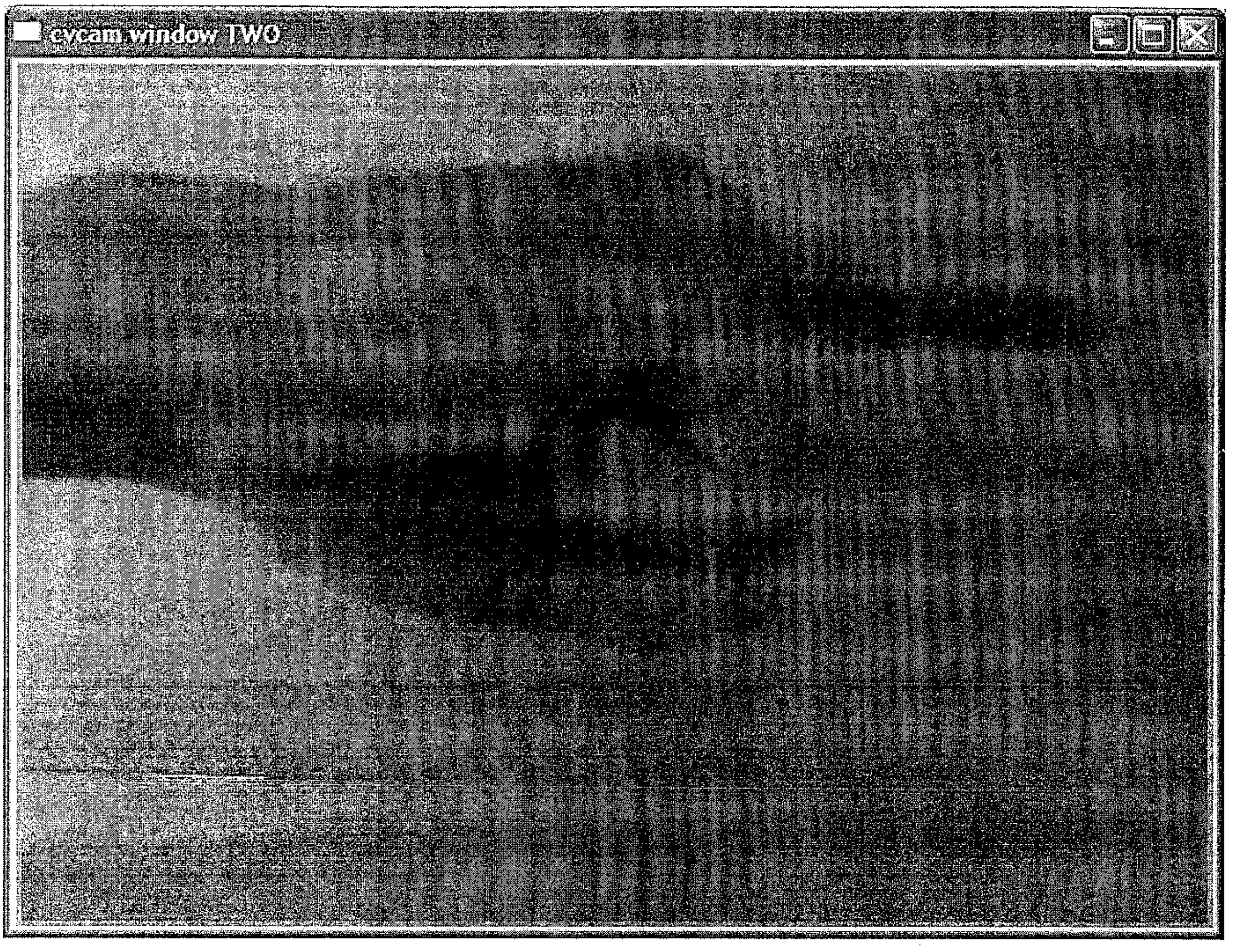



Image 2: Image Captured from XZ Plane

Both the images YZ Plane and XZ Plane is produced by both the webcams captured simultaneously from different orientation with each camera is 90 degree apart from each other. This unique orientation making sure that the device can capture the hand in 2 different perspectives to get 2 different planes of views.

This hand is captured inside an enclosed environment in pink to provide a uniform background to ease the foreground and background segmentation. The enclosed case make the whole environment to be in dimmed pink light condition which make the webcams to be operated in low light mode.

The webcams produced a not promising quality of images with a lot of noise and it is predicted to be affecting the result of foreground and background segmenting.

Overall the milestone of image capture is reached with both cameras able to supply images simultaneously but the frame rate is reduced from original 30fps to 20-22 fps. One of the reasons of the reduction of frame rate is the limited bandwidth of USB which connects both the cameras with the computer. However, the webcams are connected to the computer through the USB must not share the same hub or else, the frame rate will further drop.

## 9.2 Image Processing Module Result

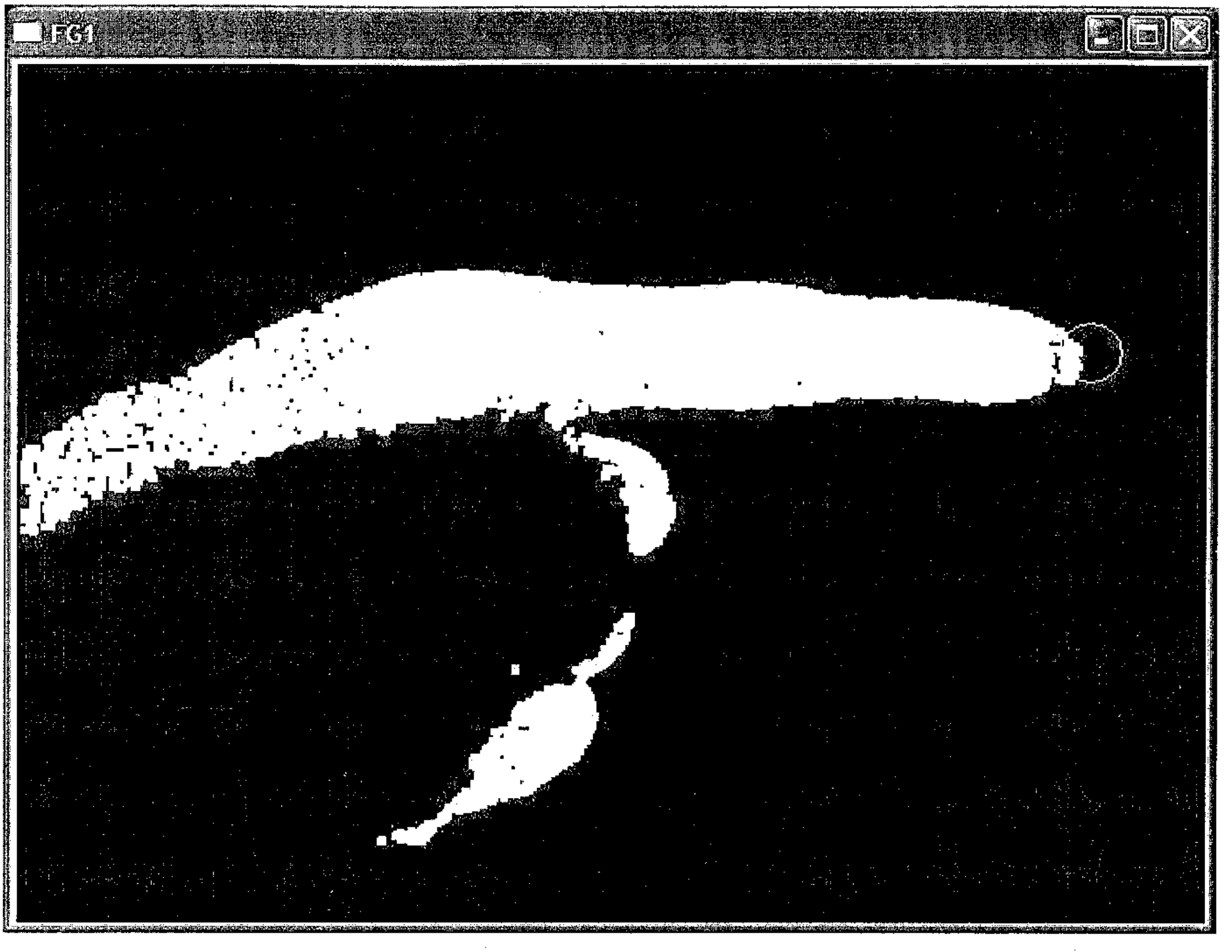



Image 3: Foreground Segmented YZ Plane

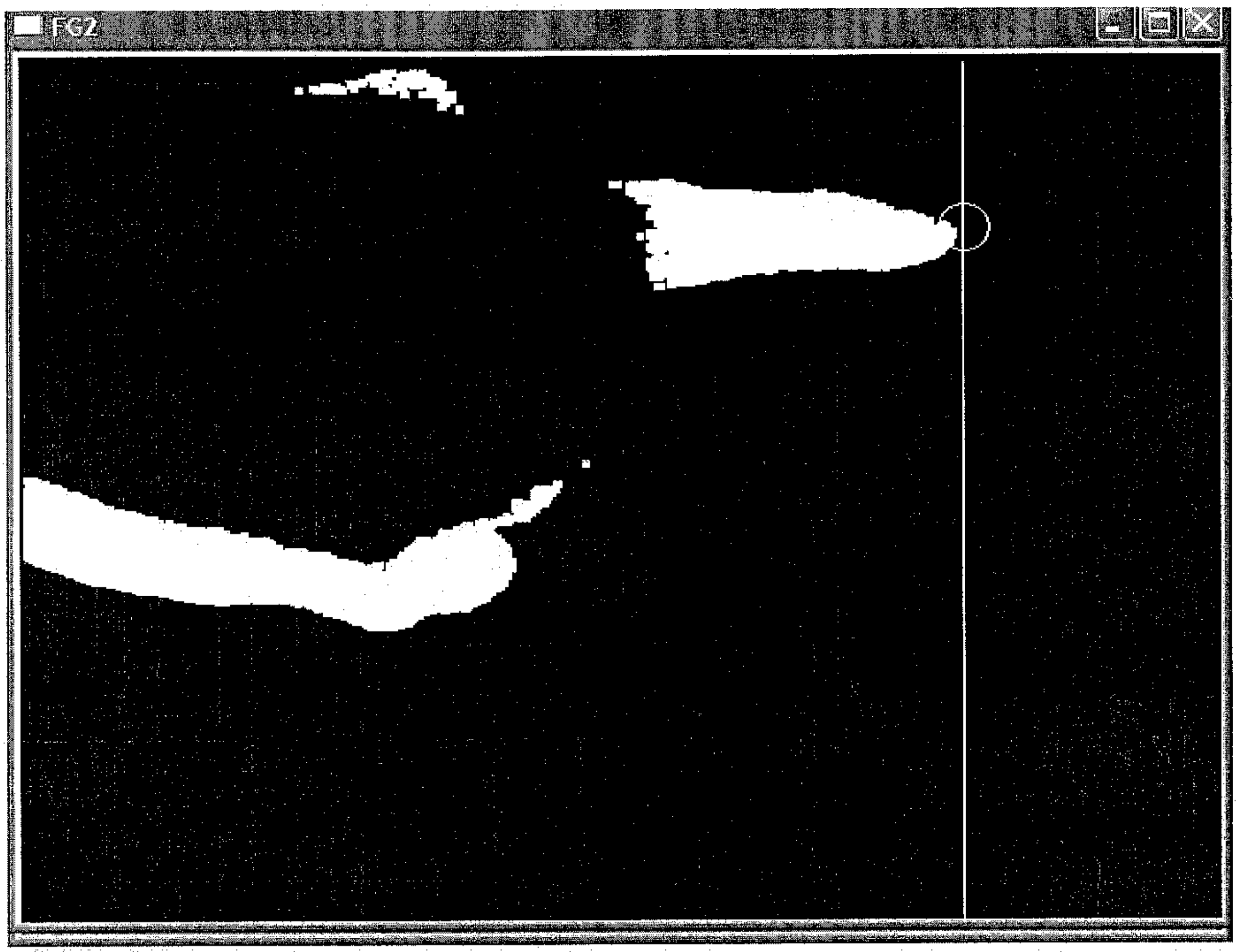



Image 4: Foreground Segmented XZ Plane

There are 3 main difference of the images compared to the original images captured from the image capture module after the images going through the image processing module. Firstly, the white blob in the images represents the foreground segmented from the background calculated. To reduce the noise before the peak detection, the foreground is eroded and diluted before it enters the peak detection function. The process of erode and dilution deleted the small pixel than fall below threshold and then increase back the foreground to original size by dilution of pixels.

After that, the foreground enters the process of peak detection. The CPU will calculate the nearest pixel starting from the right hand side row by row with the loop function. After that it will place a circle on the detected peak to indicate the peak to the user. Each peak from each plane, will then store inside a variable for coordinate matching module.

The coordinate matching module will take both the 2D coordinate and then calculate the 3D coordinate based on the reference axis which is the common Z axis. The module will draw a line inside the XZ Plane to show how close the difference of the matching Z axis with one another. The closer it is, the accurate it is of the 3D coordinate.

## 9.3 Image Display Module Result

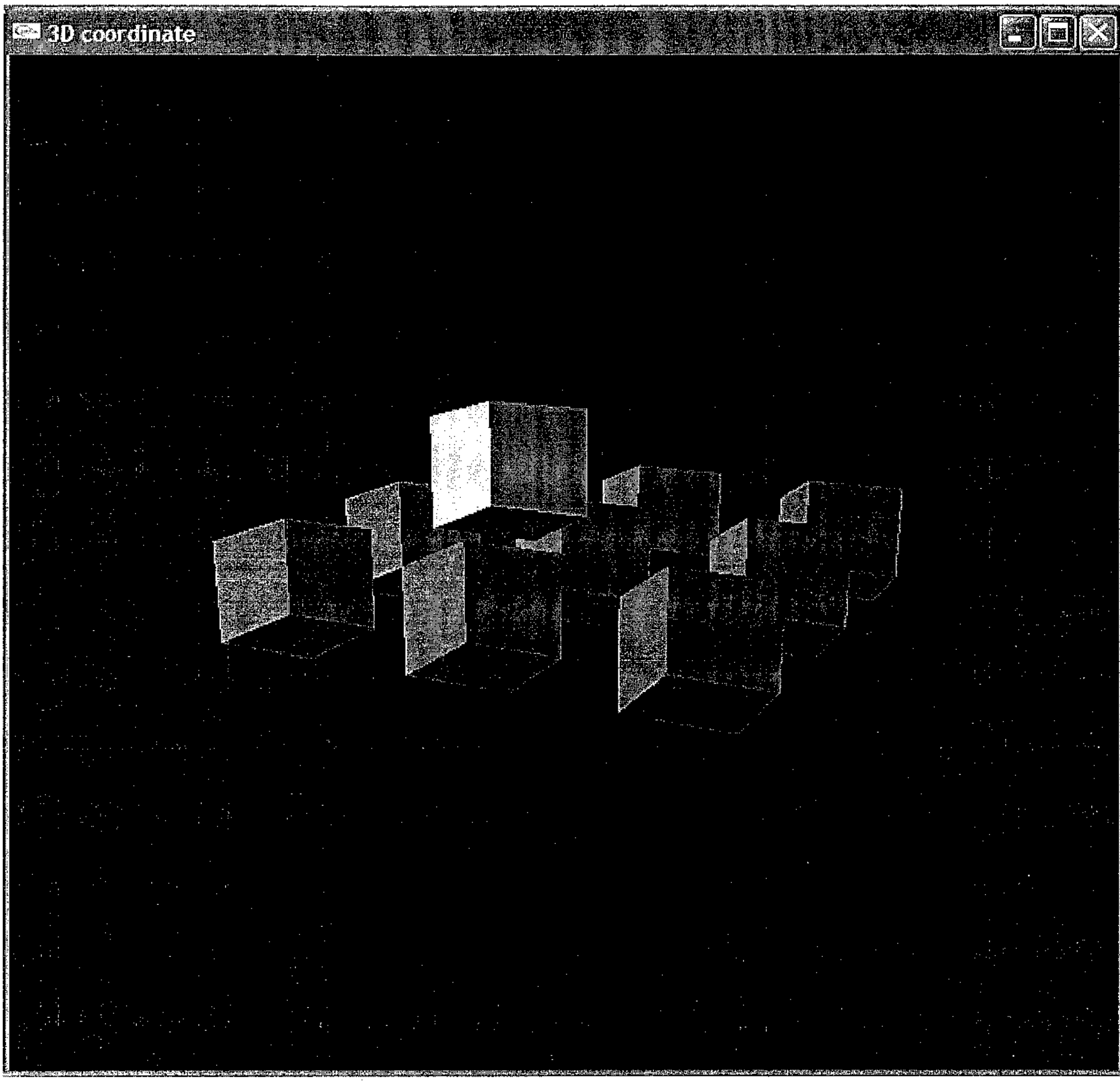



Image 5: 3D Cubes for Image Display

This image display module is a simple application with 9 red boxes lying flat to represent the orientation of a key pad. The green box is the cursor that is drawn on the coordinate supplied from the image processing module. The green boxes will move around in the spaces and the red boxes act as the indication for the place of the green box whether it is further away from the user or closer to the user.

Before entering this image display module, the 3D coordinate supplied from the image processing module is being adjusted to a factor because the detected area is not similar to the 3D environment. When this 3D Spatial Sensor is ported out for other application, the same factoring function needs to be done before it is shown in on the screen.

# 10. Testing

This prototype went through 2 types of testing to ensure that each milestone meet the requirement before stepping into the next stage. Each milestone is very important so that the further stage can be correctly implemented with minimum error and the overall project and meet the requirements and objectives. The testing phase as indicate in the schedule includes both module testing and overall system testing. Both the testing is black box testing where a test case is presented so that the minimum function is archive with the input.

### 10.1 Module Testing

Image Capture Module

| **Test Case** | **Result** |
|---|---|
| **Simultaneously capture 2 image frames** | Yes |
| **Frame that hold 2 webcams 90 degree apart** | Yes |
| **Maintaining 30 fps** | 20-22 fps |
| **Showing both images on the screen** | Yes |
| **Supply common images format to image processing module** | IPLimage is used for further processing |

Table 1: Image Capture Module Test Case

Image Processing Module

| **Test Case** | **Result** |
|---|---|
| **Background and Foreground Segmentation** | Yes with minimum noise |
| **Erode Images** | Yes |
| **Dilute Images** | Yes |
| **Showing both images on the screen** | Yes |
| **Detect Peak** | 60% Detected correctly |

Table 2: Image Processing Module Test Case

Image Capture Module

| **Test Case** | **Result** |
|---|---|
| **Display 3D environment** | Yes |
| **Refreshing the frames** | Yes |
| **Cursor that are shown based on 3D coordinate** | Green Box |
| **Factor the 3D coordinate according to the environment** | Yes |

Table 3: Image Display Module Test Case

## 10.2 System Testing

System testing is the testing on overall functionalities whether does this device meet the project requirement proposed. After all modules testing, the component functionalities are able to get pass with some component able to archive minimum requirement.

3D Spatial Sensor achieve the main objective to develop a user interface that capture the finger tips and generate a 3D coordinate without the marker. The whole process makes use of image processing and peak detection with 2D coordinate matching. The output of the device is not as accurate but it proves the concept of this idea is able to produce 3D coordinate that let user interface with 3D environment inside computer.

Another main objective achieved is the coordinate is can be used and drawn in the 3D environment with the help of the factor function. It proves that application can be built based on this setup. It is proven by testing and viewing the simple sample 3D application which consists of 9 boxes and a cursor box, the cursor box is able to move according to the finger movement in the detected area.

# 11 Recommendation

## 11.1 Image Acquisition

Image acquisition is the early part of the whole application. At the current moment, both the hardware and software of the prototype has a limitation which compromise the accuracy of the prototype. Low image quality that produced by both the webcam resulted the image processing module to wrongly calculate the output of the result. While the image capture function that grabs image frames from the webcams need to be improved so that the frame rate of the webcams do not reduce which will decrease the reaction of the resultant coordinates.

### 11.1.1 Multi-processing

Image capture part is the part that retrieves image frames from the webcams and supplies them to the image processing module. In this prototype, the image capture part is a simple image capture function that retrieves images from the webcams sequentially making the frame rate to be halved. The lower the frame rate, the lower the responsiveness of the result and making the image display module seems to be lagging.

To further improve this prototype, it is recommended to develop an image capture module that utilise the ability of multi-processing of the current processor. With multiple threads to be implemented for all simultaneous function of this prototype, the result of this product will be very fast and responsive.

### 11.1.2 Better Quality Camera

A better quality camera will affect the result of this prototype a lot. With the higher resolution of each camera, the detected finger tips will be more accurate because a slight movement of the finger on the bigger image frames will change the result. A higher resolution camera to capture the finger movement is essential for critical application to supply minute changes of 3D coordinate.

Some commercial camera does supply bigger field of view with bigger CMOS / CCD sensor and wide angle lens. Bigger field of view means bigger overlapping region of both cameras. Bigger overlapping region will give a bigger detected area to this 3D Spatial Sensor. Bigger detected area has a main advantage where the user has a lot of spaces to move around and more icon / events can be accommodate into the detected area for the user to touch.

Choosing a CCD image sensor camera compare to a CMOS image sensor camera will help in reducing the noise of image frames supply to the image processing module because CCD camera produce a higher quality image than CMOS camera with a higher resolution.

## 11.2 Image Processing

Image processing play a main role in this project by analysing the image capture, segment out the foreground from the background, detect the finger tips and calculate the final output as 3D coordinates. In this prototype, the inaccuracy of the outcome came from this simple image processing module. Before commercializing it to match the standard of the market, the image processing module must be enhance to improve the accuracy and the responsiveness.

### 11.2.1 Advanced Background and Foreground Segmentation

A good background and foreground segmentation process will produce a crisper foreground and better finger tips. There are lot background and foreground segmentations with different techniques and approach to segment out the foreground from the background. In this prototype, a simple background and foreground segmentation does not promise a good result and resource hungry. There are more suitable background and foreground segmentations for 3D Spatial Sensor with minimum resource but produce better result for finger tips recognition.

### 11.2.2 Improved Finger Tips Detection

Currently the finger tips detection is limited to detecting the highest peak of the foreground which eventually is the finger tips. But this method has a lot of flaws including it only limited to single point detection which left no room for multiple touch spatial sensors. An improved finger tips detection needed to be implemented so that it can accommodate multiple touch and accurately detect the place of the finger point. Another advantage of finger tips detection is the coordinate matching module will be more accurate and do not always match and produce a false result.

## 11.3 Image Display

Image display part of this image processing project can also be associated to the application part of the sensor developed for this project. The prototype of this project consists of an image display part which is represented by several boxes in 3 Dimension and another box in green colour represents the pointer. The 3D coordinate resulted from the image processing will redraw the green coloured box each time a new 3D coordinate is generated. The current prototype of image display part is just to show how the 3D coordinate is inputted inside the 3 dimension world and we can arrange the green box by moving our finger directing the green box.

### 11.3.1 Pluggable Front-End Application

This part of the project can be upgrade to pluggable and exchangeable with different type of application to suit other need. The application can be extended to the proposed application like keyboard simulation, 3D gesture recognition and 3D graphic modelling. When it is needed to change to another application, a pluggable format of file can be install inside into the application to perform another function. The user can choose to load which function they intended to use and switch to another function on-the-fly.

This pluggable front-end application is feasible to be developed as the input to the plug-in is just 3D coordinates supplied from processing the images.

### 11.3.2 Application Programming Interface (API)

Since the software is straight forward, redeployable and modular in the development phase, this project can be upgraded to API where the whole source code can be in the form of a library and the functions can be called for other purpose. This API can help other programmer to develop more 3D application that can associate to this 3D Spatial Sensor.

# 12. Conclusion

The prototyping process will not only help us to visualize the progress while understanding the functions of spatial sensor, it will ease the development by completing step by step progress based of level of difficulties.

The development of 3D spatial sensors require a lot of efforts, resources and expertises from different area such as computer language programming, computer vision, image processing and mathematical theory. But, developing a 3D spatial sensor will enhance human experience in human computer interaction and create a lot of opportunities for application development such as input devices and virtual reality applications.

Upon finishing this device, a patent can be filed so that it will recognize this device is a novel and unique idea whereas this project demonstrate the functionalities and how feasible is this sensors.

# 13. References

1. Sharp, Jeffrey L.(2004). *U.S. Patent No.6856259*. Newark, CA: U.S. Patent and Trademark Office.

2. Rubbert, Rudger (1999). *U.S. Patent No.6359680*. Berlin, DE: U.S. Patent and Trademark Office.

3. Morrison, Gerald D. et. Al (1999). *U.S. Patent No.*6947032. Calgary, CA: U.S. Patent and Trademark Office.

4. Xuefeng LIANG, *Nonmember and* Tetsuo ASANO, A Linear Time Algorithm for Binary Fingerprint Image Denoising Using Distance Transform

5. Xiaoqian Jiang, Wanhong Xu. Carnigie Mellon University, Contactless hand recognition

6. Weiying CHEN, Ryuji FUJIKI, Daisaku ARITA and Rin-ichiro TANIGUCHI, Kyushu University (2007). Real-time 3D Hand Shape Estimation using Multiple Cameras

7. Codeproject : ICE_WIZARD, How to Select an Object in OpenGL <http://www.codeproject.com/KB/openGL/openglselectobject.aspx>

# Appendix

## Conceptual Diagram

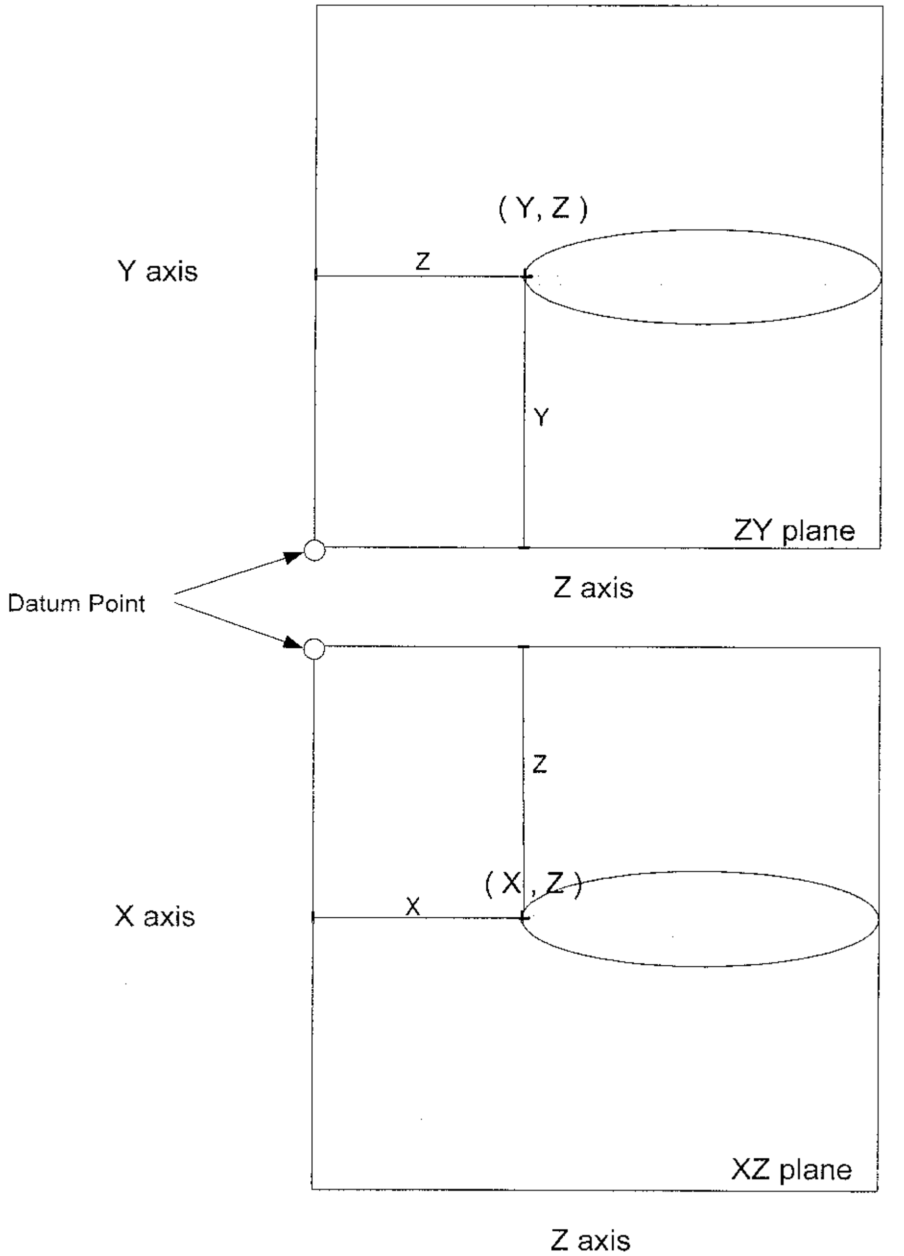



Figure 1: Single Touch Diagram

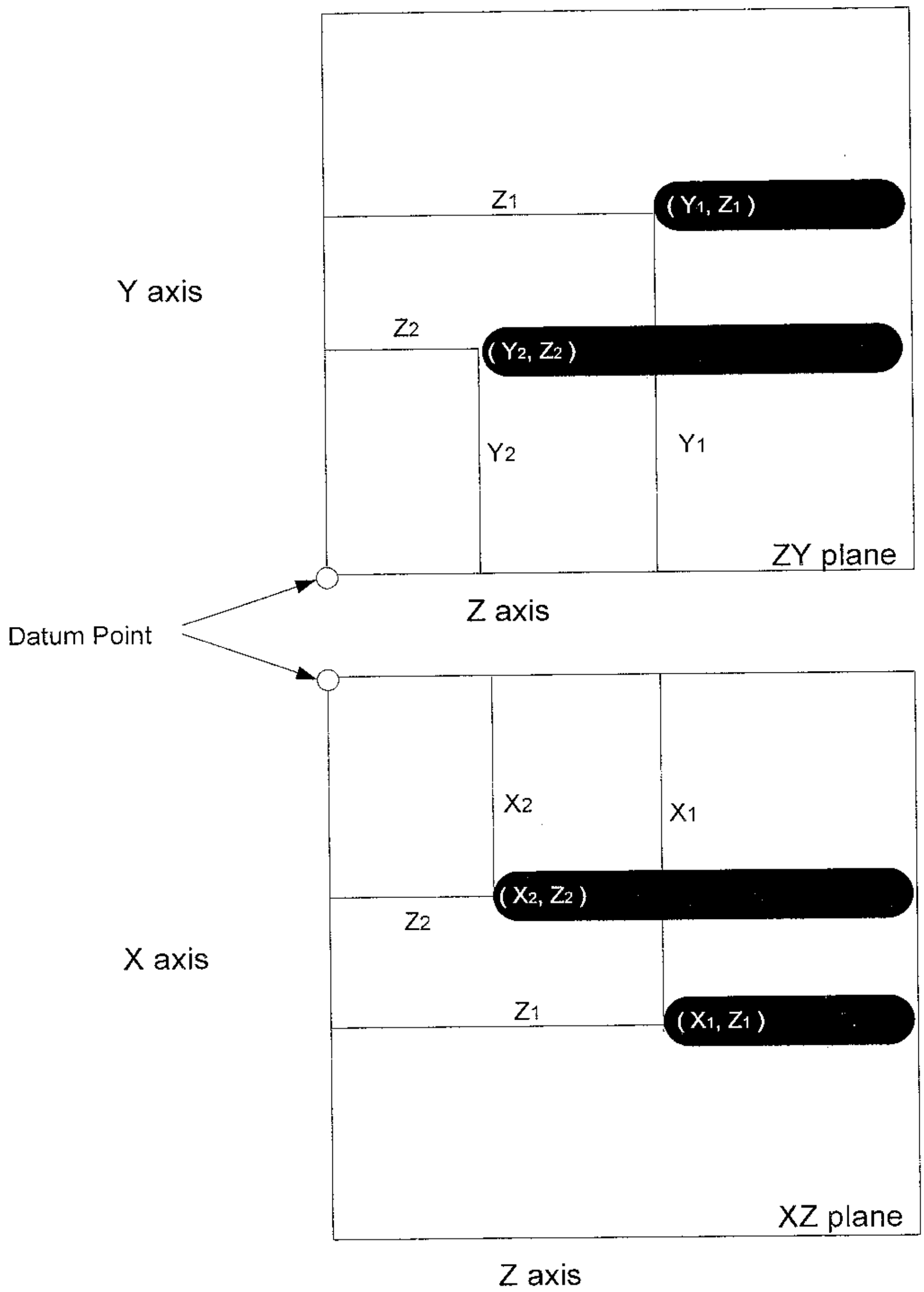



Figure 2: Double Touch Diagram

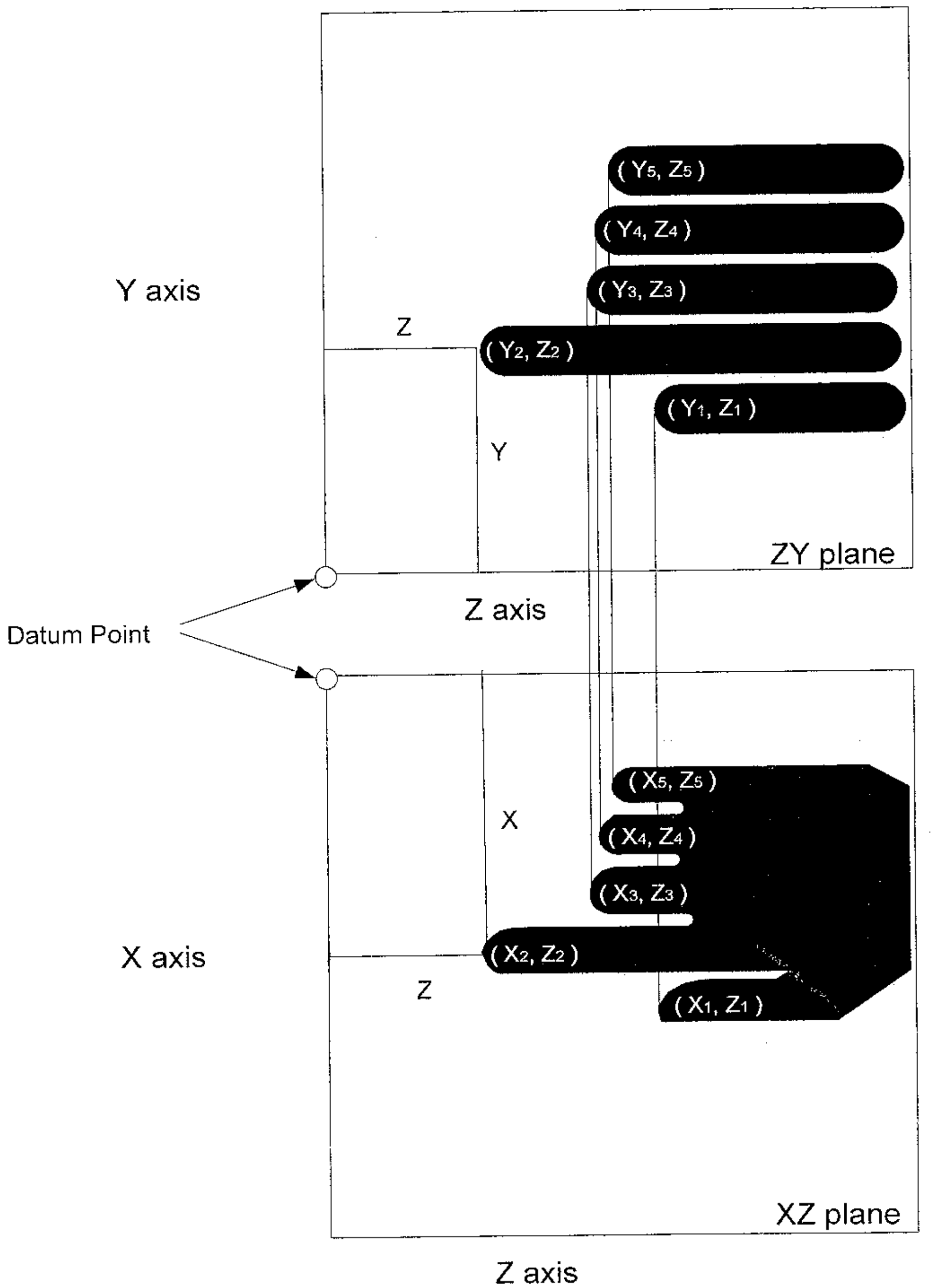



Figure 3: Multiple Touch Diagram