適用於內嵌光感測器顯示器之多使用者及多點互動系統

碩士研究生：董軒宇　指導教授：謝漢萍教授

黃乙白副教授

國立交通大學電機學院　　光電工程研究所

摘　要

隨著互動技術的不斷演進，人機互動的方式已經從平面上的觸碰躍升至空間上的互動。三維互動系統不僅可以提供更加人性化的操作方式，更可令使用者與立體影像的互動更有真實感。然而，隨著可攜式電子產品在生活上的重要性不斷提升，至今卻仍未發展出適用在可攜式電子產品上的三維互動系統。為了滿足可攜式電子產品在可攜性以及互動範圍上的要求，使用結合內嵌式光感測器之顯示器以及外加光筆的互動系統會是相當不錯的選擇。

在這樣的架構下，本論文提出三種不同的三維互動系統以滿足不同的需求。T Mark System 利用貼有 T 型遮罩的光筆來進行互動。藉由分析內嵌式光學感測器所捕捉到的影像，光筆的三維座標(X, Y, and Z)、旋轉與傾斜的角度(θ and φ)都可以被計算出來。為了辨識更複雜的動作，本論文提出了兩種不同的用途的三維多點互動系統。Color Filter Based 利用可見光之光筆以及顯示器既有的彩色濾光片來判別出個別光筆的三維位置。此架構在影像嚴重重疊時亦能運作，故此架構適合使用於單使用者的多點觸碰。然而使用可見光做為互動光源會降低顯示器的影像品質。為了解決這個議題，Multi-Mark Based 利用紅外光取代可見光來做為互動光源。藉由在不同的光源上貼上不同的遮罩並分析所擷取到的影像，亦可以獲得不同光筆的三維位置。然而此系統在分析重疊影像的能力較不足，此系統較適合使用於多位使用者的單點互動。

Multi-user and Multi-touch Systems for Display With Embedded Optical Sensor

Student: Hsuan-Yu Tung Advisors: Dr. Han-Ping D. Shieh
Dr. Yi-Pai Huang
Institute of Electro-Optical Engineering
National Chiao Tung University

Abstract

In recent years, interactive systems have been evolving from 2D touch into 3D interaction. 3D interactive systems can not only provide intuitive operation but also sense of reality when users interacting with 3D images. However, little research has been done on the 3D interactive systems applied on portable devices. For satisfying the requests of portable devices, 3D interactive systems combined embedded optical sensor and additional light sources will be a better solution.

In this structure, we propose three kinds of 3D interactive systems for different requirements. T Mark System uses light source pasted with T mark to detect the 3D positions (X, Y, and Z)、rotation angle (θ) and tilt angle (φ) of single user. For detecting more complex actions, two multi-interactive systems will be proposed. Color Filter Based uses visible light sources and color filters to distinguish different light sources and calculating the 3D positions of each light source. This system is still workable when the projections of light sources are seriously overlapping. However, the visible light sources will reduce the image quality of display. For solving this issue, Multi-Mark Based replaces the visible light sources by IR light sources pasted with different marks. After analyzing the projections of these light sources, the 3D positions of each light source can be obtained. However, Multi-Mark Based cannot detect certain overlapping conditions perfectly. This system is suitable for multi-user.

誌　謝

首先向指導老師謝漢萍教授及黃乙白副教授至上最誠摯的感謝。在碩士這兩年的求學過程中，實驗室了提供了相當自由且愉快的研究氣氛，讓我得以體會做研究所需要的積極與熱情。老師對於研究態度及英文能力的指導，更令我意識到自己的不足進而努力的自我提升。在如此豐富的研究資源以及完善的研究環境之下，此篇論文得以完成。此外，特別感謝各位口試委員所提供的寶貴意見，讓此篇論文得以更加完備。

在這兩年的研究時光裡，特別感謝國振學長的熱心指導。其淵博的學識往往在我盲目摸索時給予我一盞明燈，並且時時修正我的研究方向免於落入象牙塔中。更重要的是，學長對於自我的嚴格要求以及對事物的勇於嘗試提供了我一個在人生道路上良好的學習典範。另外我要感謝友達光電諸位合作同仁在我研究上的指導與建議，並且提供我實務上的經驗讓實驗得以順利進展。

在實驗室度過的這些日子裡，芳正學長、凌嶢學長、育誠學長、精益學長、奕智學長、柏全學長、韻竹學姐、志宏學長給我的照顧與指導，讓我得以在研究之路上走得更加順遂。另外我要特別感謝致維學長時常在深夜裡陪著我在實驗室繼續努力，他那爽朗且親切的笑聲在我腦中留下不可磨滅的印象。

與同學們，思頤、博詮、志堯、博元、禹辰、立偉、又儀以及昌毅一起努力奮鬥的過程，更是我珍貴的回憶。此外也很感謝張雅惠、李穎佳和溫純敏三位助理小姐以及實驗室的學弟妹們，讓實驗室裡時常充滿著歡笑。此外，也謝謝跟我同組的明青跟書怡學妹，儘管我這麼不成才還是非常親切的給予我建議及鼓勵。

最後我要感謝我的家人以及我那位尚未出現的女朋友，謝謝你們對我的包容與支持令我可以無後顧之憂的努力做研究，並得以完成此份論文。謝謝各位，在此致上萬分感謝。

Table of Contents

Figure Captions

–

Table Captions

Chapter 1

Introduction

1.1 Preface

In recent years, there has been a dramatic proliferation of the research concerned with interactive technology. The applications of two-dimension (2D) touch technique have been dramatic increase due to the intuitive user interface and the supporting of multi-touch. In the mobile systems, touch interface even substantially replace the traditional keyboard interface. However, with the rise of Wii Remote of Nintendo [1][2][3] (as shown in Fig. 1- 1) and Kinect of Xbox 360 (as shown in Fig. 1- 2), a lot of research have be done[3][4]. The applications of three-dimensions (3D) interactive technology have been a tremendous wave.

Fig. 1- 1 Wii and Wii Remote [4]

Fig. 1- 2 Xbox 360 and Kinect [5]

3D interactive system can provide more information than 2D touch system. By the way, 3D interactive system can distinguish more complex actions like fishing、pushing or even catching. In the field of entertainment、communication or education, 3D interactive system can support not only immersive operating experience but also more friendly interface. Moreover, if we combine 3D interactive system with 3D display, we can interact with 3D image more truly. Thus, 3D interactive function is very important for future life.

1.2 3D Interactive System

In general, 2D interactive system can recognize the 2D positions (X and Y) of fingers when user contacts with panel [6][7][8]. However, touching the panel is not allowed in some places likes in operating room of hospital. On the other hand, 2D interactive system cannot provide sufficient degrees of freedom for executing complex actions. Moreover, with the rapid development of 3D display, a new interactive system which can help us to interact with 3D images had become more and more important. For these requirements, 3D interactive systems which can detect the 3D positions (X, Y and Z) of user need to be designed.

However, 3D interactive system needs to detect not only the 2D coordinates but also the depth information (distance between user and panel). For this goal, some different detection technologies are necessary[9][10][11]. With the different detection devices, 3D interactive system can be divided into Machinery Based、 Optical Measure Based and Optical Sensor Based. The schematic overview of 3D interactive system is shown in Fig. 1- 3.

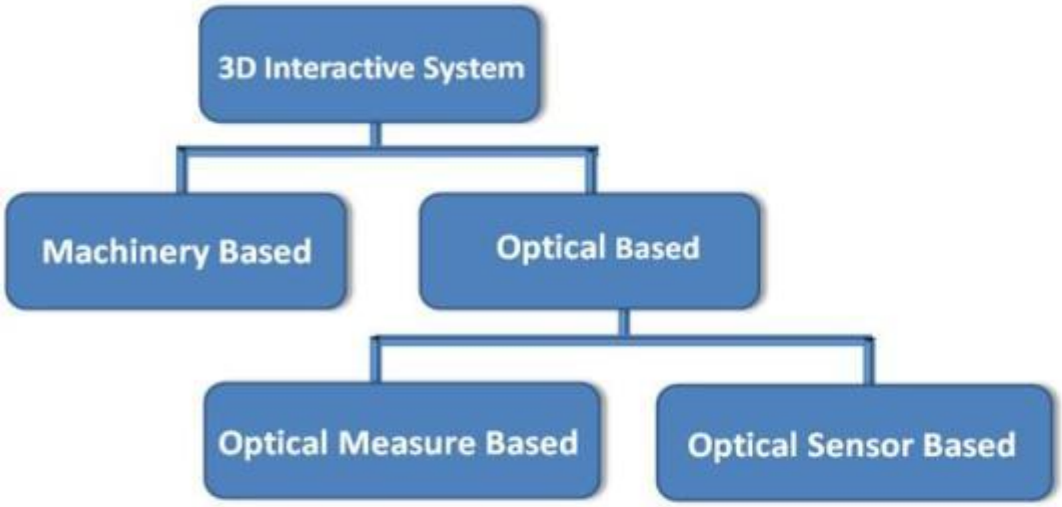


Fig. 1- 3 Classification of 3D interactive systems

Machinery Based

Machinery Based (as shown in Fig. 1- 4) uses additional devices (accelerometer、gyro and so on) to detect the information of three dimensional positions. Although these additional devices can help us to obtain more accuracy 3D positions, but the heavy and expensive devices will be issues for portable devices.

Fig. 1- 4 Machinery Based 3D interactive system [12]

Optical Based

Optical Based can be divided into two major categories by different detection methods：Optical measure based (as shown in Fig. 1- 5) and optical sensor based (as shown in Fig. 1- 6). Optical measure based uses some properties of light to calculate the 3D positions of user directly. Optical sensor based uses optical sensor (camera or embedded optical sensor) to capture image of user, and then uses some skills of image processing to obtain the position information.

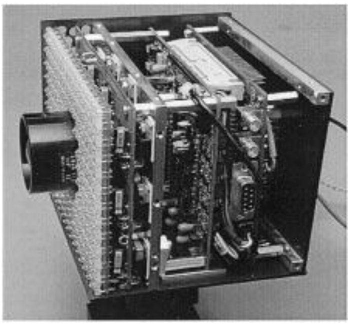


Fig. 1- 5 Time of Flight camera [13] Fig. 1- 6 Continuous interaction spaces [14]

1.3 Motivation and Objective

Nowadays, 3D interactive system has been widely used in different fields. No matter in the areas of entertainment、communication or education, 3D interactive system can bring us unprecedented experiment. We can choose appropriate system for different demands. However, there's still no adequate 3D interactive system for portable devices. In order to satisfy the increasing demand, developing 3D interactive systems which are suitable for portable devices have become urgent priority.

When using in portable device, the heavy and expensive devices of Machinery Based will let user feel inconvenient. Optical measure based can be combined with portable devices, but it usually needs large operating distance because the velocity of light is very fast. By the way, optical measure based is not suitable for portable devices too. In optical sensor based, camera is widely used to capture image of user. Due to the high resolution of CCD (charge-coupled device), camera can recognize complex actions. However, blind range (as shown in Fig. 1- 7) will happen when users operate near the cameras. This issue will cause the discontinuous when user switches the operating mode between 2D touch and 3D interaction. Moreover, camera based usually has large volume. These features will be disadvantages when using in portable devices.

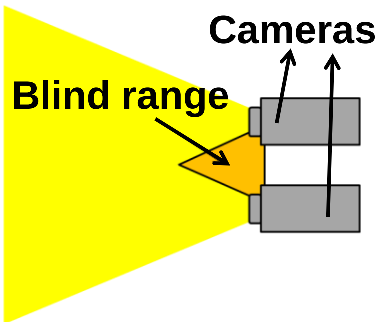


Fig. 1- 7 Blind range of camera

An adequate solution to the application of portable devices is the embedded optical sensor (as shown in Fig. 1- 8). Due to the optical sensors are embedded with display, the issue of blind range can be eliminated. By the way, the volume of system can be reduced too. For lower cost and using in portable device, optical sensor based is a better choice to be our platform. However, the resolution of embedded optical sensor is not good as camera.

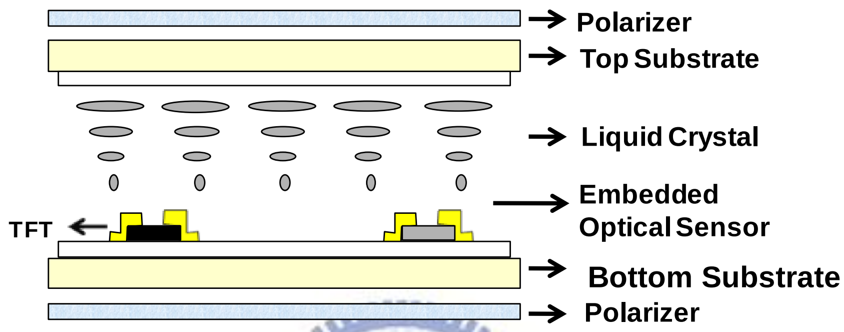


Fig. 1- 8 Display with embedded optical sensor [15]

A sound 3D interactive system should support the multi-interactive function. Multi-interaction can support more freedoms to interact. By the way, some complex actions like fishing、pushing or even catching can be executed. 3D interactive systems can be achieved by using several light sources to interact with embedded optical sensor. These systems will be proposed in this thesis.

1.4 Organization of This Thesis

The organization of this thesis is as following: Various kinds of relational 3D interactive systems will described clearly in **Chapter 2.** The principles、advantages and drawbacks of each 3D interactive system will be discussed. The motivation and objective of my thesis will be illustrated in this chapter. A single 3D interactive system called T Mark System will be introduced in **Chapter 3**. This system provides more degree of freedoms than previous systems. The skills of image processing using in this system will be the foundations for next proposed 3D interactive system. In **Chapter 4**, a 3D multi-interactive system called Color Filter Based will be presented. This system uses embedded optical sensor to capture the images of user. Color filters and additional visible light sources were used to distinguish multi-touch or multi-user. The principle, algorithm and experimental results will be unraveled. However, the visible light sources will reduce the display image quality. For this issue, another multi-interactive system called Multi-Mark Based will be proposed in **Chapter 5**. Multi-Mark Based uses infrared light sources to instead the visible light sources. Several different marks were combined with different light sources. By the way, the projection of each light source will be significantly different. Projection of each light source can be obtained by some skill of image processing and get the 3D positions of each light source by analyzing the projection images. Finally, some conclusions and future works will be shown in **Chapter 6**.

Chapter 2

Prior Arts of 3D Interactive Systems

Several representative technologies of 3D interactive system will be introduced. These systems provided us some ideas to establish 3D interactive systems for portable devices.

2.1 Classification of 3D Interactive Systems

In recent years, 3D interactive systems have become more diversity and user friendly. 3D interactive systems can be divided into two major types: Machinery Based and Optical Based. Although these systems use different detection methods, both of them focus on how to detect accuracy position information and distinguish the action of user. In the chapter, the principles and some impressive research of Machinery Based and Optical Based will be described in detail.

2.2 Machinery Based

Machinery Based is the first appeared 3D interactive system. In Machinery Based, users were asked to wear additional devices for detecting 3D positions (X, Y and Z) or actions. Some devices like gyroscope which can measure orientation based on the principle of conservation of angular momentum. The other devices like accelerometer which can measure the acceleration of device. By detecting the positions continuously, actions of user can be obtained. Data glove (as shown in Fig. 2- 1) contains facilitates tactile sensing and fine-motion control so that it can be used to calculate gesture of user accurately.

Fig. 2- 1 Data glove [16]

These devices will collect data of user and sent back to computer to recognize the intentions of user by comparing these data with database. Finally, 3D interaction between users and system can be achieved. Due to these devices had developed for many years and the accuracy of detection results is good, Machinery Based can be used for precise works likes painting or medical treatment. However, these devices always are heavy and expensive due to their complex structures. By the way, Machinery Based is not suitable for using on portable devices. The following are some impressive 3D interactive systems in Machinery Based.

2.2.1 Data Glove

Data glove had been developing for a long time. Data gloves are gloves contained different kinds of sensors to detect movement of user. Data gloves can be divided into three majors by different detection devices：Optical glove (as shown in Fig. 2- 2) 、strain gauge glove (as shown in Fig. 2- 3) and mechanical glove (as shown in Fig. 2- 4). Optical glove puts fiber on each finger and sets the optical sensor on the top of each finger. When the finger is straight, the light intensity detected by optical sensor will be strong. When the finger is bent, the intensity of detection light will be weak. Strain gauge glove uses the conducted ink or strain gauge to detect the degree of bending of fingers. The voltage of these material will change when their shape change. We can judge the degree of bending by the change of voltage. The third one is mechanical glove. Linkages are established on the joints of each finger to detect the status of fingers. When the fingers are bent, the degree of linkages will be larger. Combining with some positioning devices, data gloves can provide 6 degrees of tracking (X, Y, Z, Yaw, Pitch and Roll).

Fig. 2- 2 Optical glove [17]

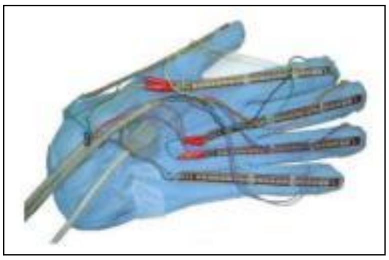


Fig. 2- 3 Strain gauge glove [17]

Fig. 2- 4 Mechanical glove [17]

2.2.2 Wii

Wii is the seventh generation game console launched by Nintendo in 2006. Quickly, Wii had become a popular game console because of its special controlled method. The controller called Wii Remote can detect the 3D positions and distinguish the action of user. With the special controller, Wii can provide immersive gaming experience to users. Users can use physical movement to be the controller. The key detection components of Wii were three coordinates accelerometer (as shown in Fig. 2- 5) and infrared CMOS sensor (as shown in Fig. 2- 6) and sensor bar.

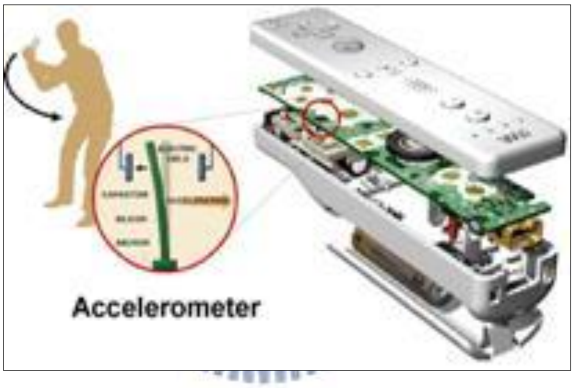


Fig. 2- 5 Three coordinates accelerometer [18]

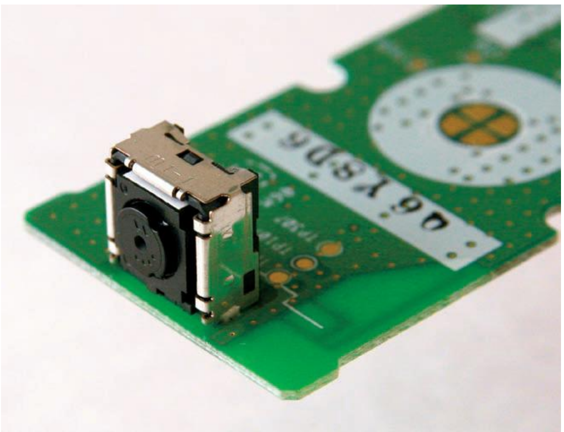


Fig. 2- 6 Infrared CMOS sensor [18]

Three coordinates accelerometer can detect the acceleration along three axes (X, Y and Z) by measuring the spacing of plat capacitor. In plate capacitor, the capacitance can be described as Eq. 2-1. The capacitance is related to the electric charge Q and the voltage V. A is the area of each plate、k is dielectric constant and d is the distance between two plates.

$$C = \frac{Q}{V} = \frac{A}{2\pi kd} \qquad \text{Eq. 2- 1}$$

Form above equation we can find that the capacitance is inversely proportional to the distance between two plates. We can detect the acceleration by measuring the variation of capacitance. Wii use three orthogonal capacitance accelerometers. The plates of each capacitance were movable. By the way, when the capacitances changed, we can calculate the accelerations along three axes. Sensor bar contained several fan-shaped arrangements of infrared LEDs. Wii Remote uses camera to capture the image of these LEDs and calculate relational distance between user and sensor bar. Finally, above data will send to the console by Bluetooth and get the corresponding information.

Due to Wii Remote is cheaper and more common than other Machinery Based devices, there are many applications based on this system. As shown as Fig. 2- 7, if we set up the IR LED array as the light source and pasted infrared reflective film on our fingers, we can track the trajectory of the fingers. If we use IR light pen to be the interactive light source, a low-cost multi-touch electronic whiteboard (as shown in Fig. 2- 8) can be achieved. Moreover, Wii Remote even can be used to control robot (as shown in Fig. 2- 9). However, these systems cannot detect slight variation in Z direction.

Fig. 2- 7 Finger tracking system based on Wii [19]

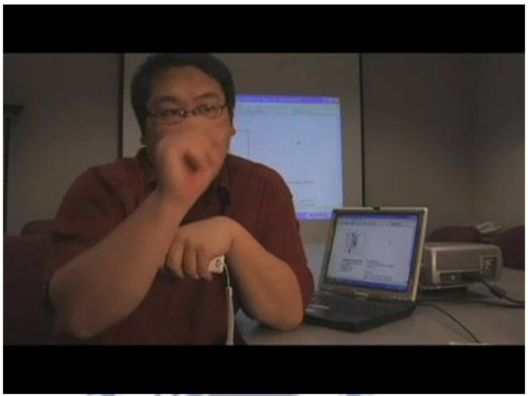


Fig. 2- 8 Electric whiteboard based on Wii [20]

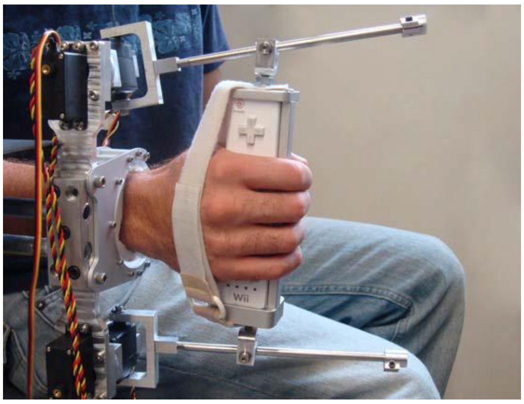


Fig. 2- 9 Robot controller based on Wii [21]

2.2.3 Other Technologies

In addition to the above technologies, there are other detection methods can help us to detect 3D positions. These technologies will be simply described as following.

Mechanical Tracker

Mechanical tracker (as shown in Fig. 2- 10) is consisted of linkages、gear wheels、sending sensors、potentiometer and encoder. The 3D positions (X, Y and Z) can be calculated by the angle change of linkage.

Inertial Tracker

Inertial tracker is designed by the principle of inertia. This system can detect the relative movement of user. Both gyro and accelerometer are inertial tracker.

Electromagnetic Tracker

Magnetic material will produce a magnetic field in free space. Electromagnetic tracker uses emitter to produce a magnetic field. The magnetic line of force will change when the magnetic receiver enters the magnetic field (as shown in Fig. 2-11). By the way, the 3D positions of receiver can be obtained.

Fig. 2- 10 Mechanical tracker

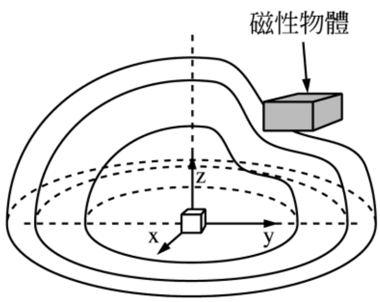


Fig. 2- 11 Electromagnetic tracker

2.3 Optical Measure Based

Optical Based is consisted of two different types：Optical measure based and optical sensor based. Optical measure based uses the properties of light to measure the position of user [15]. For example, the distance is equal to the velocity multiplied by time. If we can know the information about velocity and time, the distance between user and device can be calculated. This technique is also known as TOF (time of flight). However, due to the speed of light is very fast, TOF technique cannot apply on portable devices. The following are some impressive 3D interactive system based on optical measure based.

2.3.1 ZCam

The detection method of ZCam (as shown in Fig. 2- 12) is the Time of Flight (TOF) technology. The principle of TOF is shown as this formula：distance = time × speed. The device will emit pulsed light and records the time of flight until the receiver receive the reflected light. By the way, we can obtain the distance between device and measurement point. If full space was captured, the 3D positions of user can be measured. However, the speed of light is very fast. So far, this kind of system cannot be used in near range.

Fig. 2- 12 ZCam II of 3DV Systems Corporation [22]

2.4 Optical Sensor Based

Optical sensor based is another type in Optical Based. Optical sensor based uses optical sensor to detect the intensity of ambient light directly. By the way, the images of user can be captured. After analyzing the captured images, the 3D positions of user can be estimated. In optical sensor based, camera had been researching for a long time and getting great results. In general, camera will be used to capture image due to the high resolution of CCD (charge-coupled device). However, using camera as the sensor component will has some issues when working near panel due to the blind range.

At the first glance, embedded sensor based should be the most appropriate system for portable devices because of optical sensor is embedded in display. By the way, embedded optical sensor based can maintain small volume and eliminate the blind range issue. However, the resolution of embedded sensor was not good enough to detect the image of user clearly. The following are some impressive 3D interactive system with Optical Based.

2.4.1 Touch Light

In general, camera will be the first choice to capture images of user. The structure of Touch Light is shown as Fig. 2- 13. Touch Light is consisted of DNP HoloScreen、 two IR cameras and IR illuminant. In general, cameras will be set up around the projection screen. It will increase the volume of system and cause the blind range issue. For solving this issue, Touch Light uses HoloScreen to be the projection screen. HoloScreen is a holographic rear projection screen. It allows IR light to pass through in any direction but reflect visible light irradiating at some special angles. By the way, two IR cameras can be established behind the projection to reduce system volume and be workable near panel. Two IR cameras will obtain two images of user in different directions. After combining these two images, fused image will be obtained (as shown in Fig. 2- 14). Fused image shows the difference between two images. The depth information may be computed by binocular disparity in fused image.

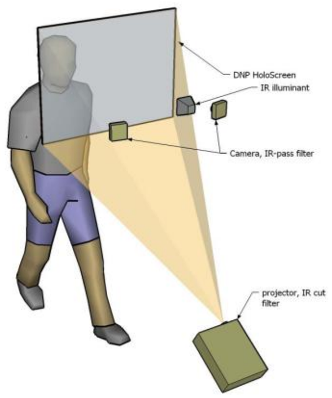


Fig. 2- 13 Structure of Touch Light [23]

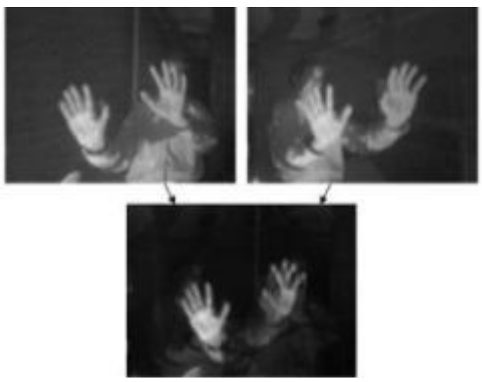


Fig. 2- 14 Fused image [23]

2.4.2 BiDi Screen

BiDi screen is a 3D interactive system proposed by MIT Media Lab. As previously mentioned, the issue of embedded optical sensor is lower resolution than camera. BiDi screen uses pinholes which are displayed by liquid crystal of display to increase the image quality of captured image. The construction of BiDi screen is shown as Fig. 2- 15. Although pinholes (as shown in Fig. 2- 16) can capture clear image clear but also reduce the brightness significantly. For this issue, MURA patent (as shown in Fig. 2- 17) was used to replace normal pinhole pattern. MURA patent not only increase the light transmission into 50% but also obtain the same image as pinhole after decoding.

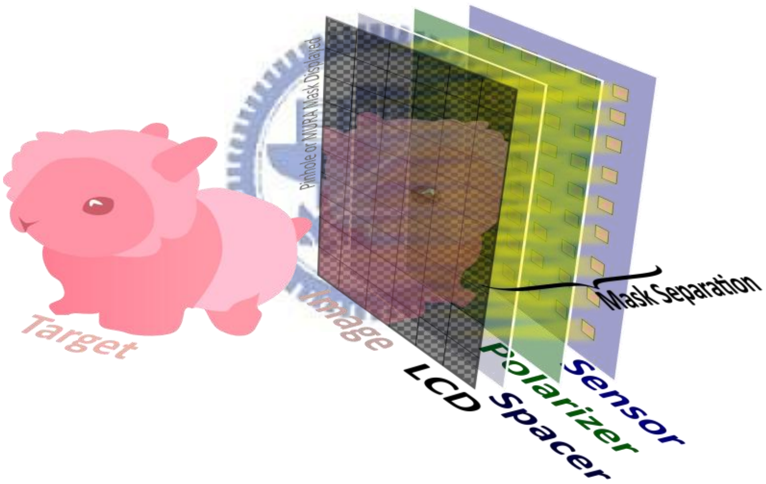


Fig. 2- 15 Structure of BiDi screen [24]

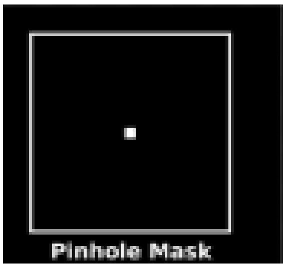


Fig. 2- 16 Pinhole [24]

Fig. 2- 17 MURA Mask [24]

Pinhole can increase the quality of captured images. After the clear image was obtained, synthetic aperture refocusing and maximum contrast operator were used to get the depth map. The flow chart can be described as Fig. 2- 18. However, although BiDi screen uses pinhole to optimize captured image, the images captured by optical sensor still were fuzzy. By the way, they still used camera to capture images of user.

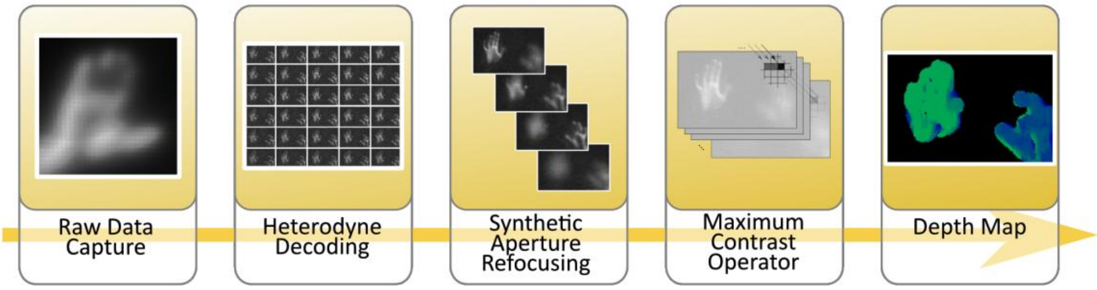


Fig. 2- 18 Flow chart of depth detection [24]

2.4.3 Kinect

In recent years, 3D interactive system has apparently moved from experimental stage attitudes to daily life. With the advent of Kinect (as shown in Fig. 2- 19) which was revealed by Microsoft in 2010, controller-free gaming mode had become a trend. It enables user to control and interact with the Xbox without any controller. User can use gestures、voice or body shaking to execute game. By the way, natural user interface was finally achieved. To achieve this amazing result, Kinect uses some interesting techniques.

Fig. 2- 19 Kinect of Xbox 360 [25]

The major components of Kinect are shown in Fig. 2- 20. It is consisted of a RGB camera、motorized tilt、multi-array microphone and a set of 3D depth sensors. At the first glance, Kinect was regarded using two cameras to judge the depth information. In fact, Kinect uses IR light source and depth camera for IR (as shown in Fig. 2- 21) to detect the depth of users rather than using two image CMOSs. If we have two separate cameras, we can use some methods of computer vision to calculate the depth information.

Fig. 2- 20 Major components of Kinect [25]

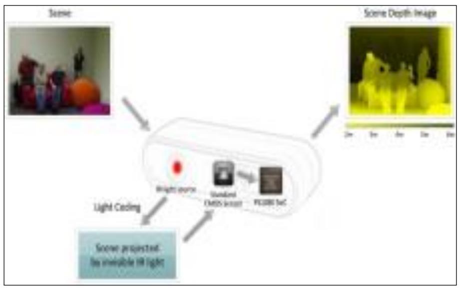


Fig. 2- 21 Depth camera [25]

The 3D depth detection method of Kinect is called Light Coding. The critical technique of Light Coding is Laser Speckle. When laser irradiates on rough surface, every position of rough surface will scatter the laser and produce random faculae [26]. The shapes of these facula are totally different everywhere. In the other words, if we record these faculae on each plane with different distance as the reference images, whole space will be marked. When user enters this space, the image of laser faculae will change. Finally, by comparing the images with reference images at the same distance, the position of user in the space will be located. The other components can also enhance the accuracy of 3D positions detection. However, Kinect still cannot detect some slight action likes motion of finger and cannot detect the positions of users near cameras.

2.4.4 Optical Sensor Based-Disparity algorithm

In addition to camera, embedded optical sensor is another choice to capture images of users. Disparity algorithm [15] is a 3D interactive system based on embedded optical sensor. The sensitivity of embedded optical sensor is not good as camera. Scanned image of user is always fuzzy and hard to calculate the 3D positions of user. For this issue, additional light source was used to enhance captured image quality. This system combines embedded optical sensor and additional lighting source is called light mode system. The structure of lighting mode system is shown as Fig. 2-22.

The additional light source can provide more position information of user. When user uses light source to control, the 3D positions of light source can be regard as the 3D positions of user's hand. The position of maximum intensity in captured image can be regarded as the 2D coordinates (X and Y). The optical sensors can be divided into L-sensor and R sensor. If object height increases, the locations of maximum intensity move away from the locations of zero object height. The Z information can be obtained by calculating disparity between two maximum intensities (as shown in Fig. 2- 23). This system can detect the 3D positions (X, Y and Z) of single user in vertical direction.

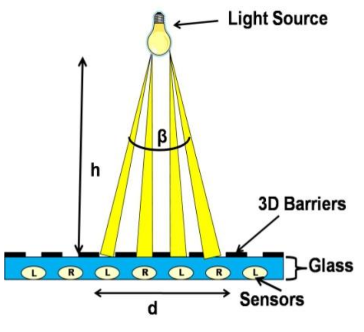


Fig. 2- 22 Structure of lighting mode [15]

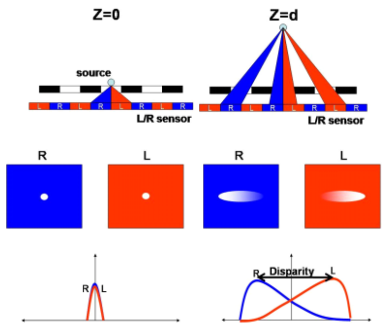


Fig. 2- 23 Principle of disparity algorithm [15]

2.4.5 Depth Detection Methods of Optical Sensor Based

Optical sensor based uses sensor to capture images of user. The 3D positions can be obtained by analyzing these images with some methods of computer vision. Some common methods will be simply described as follows.

Photometric Stereo

Photometric stereo (as shown in Fig. 2- 24) uses one fixed camera to take the pictures with different lighting conditions. After comparing these images, the gradients of surface will be solved. The 3D positions of object can be restructured by the integral of these gradients.

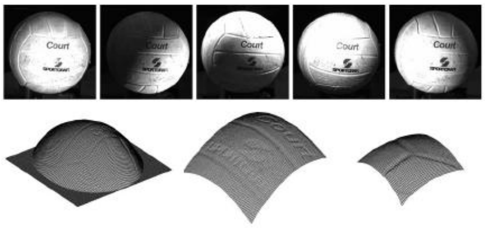


Fig. 2- 24 Photometric stereo [27]

Stereoscopic

This system uses two parallel cameras to capture images at the same time (as shown in Fig. 2- 25). Next, the corresponding points between two images can be found by multi-scale edge matching 、block matching or epipolar geometry algorithm. In the end, the depth information can be estimated by the relationship of the distance between two cameras and the focal length of each camera.

Fig. 2- 25 Stereoscopic picture [28]

Modulated Lighting

Modulated lighting system irradiates periodic sine wave on the measured object. The 3D positions of objects can be calculated by observing the brightness variations and phase variations. An example is shown as Fig. 2- 26.

Fig. 2- 26 Modulated lighting [29]

Defocus

When camera is focused on object, the contrast of image should be better than the image which was defocused. By this principle, we can compare the contrasts between defocused image and focused image to estimate the depth information. An example is shown as Fig. 2- 27. The upper image on left side is defocused image and the lower image on left side is focused image. Comparing the contrasts between these two images, the depth image (on the right side) can be calculated.

Fig. 2- 27 Defocused [30]

2.5 Summary of 3D Interactive Systems

3D interactive systems had been developing for a long time. However, all kinds of 3D interactive systems still have some issues. Machinery Based can obtain accuracy 3D positions of user but be limited by the heavy and expensive devices. This kind of system is suitable for detecting slight action.

In optical measure based, users can be released from heavy additional device. Using some properties of light, approximate position of user can be obtained. However, this system is susceptible to the external environment likes the ambient light or the surface material of object. For these reasons, this kind of 3D interactive system is suitable to use in fixed place.

In optical sensor based, camera can get better results due to the high resolution of CCD. However, cameras cannot detect the 3D positions of users near panel because of the blind range issue. It will also cause the discontinuity when user switches the operating mode form 2D touch to 3D interaction. Final, embedded optical sensor based uses embedded optical sensor to detect the ambient light. The blind range issue will be eliminated because of optical sensors are integrated with display. But the sensitivity of embedded optical sensors is not good as camera. The captured images will be fuzzy and hard to obtain the position information of users. In general, we can choose different kind of system to satisfy different requirements. For portable devices, convenience and small volume are key elements in 3D interactive system. Among different 3D interactive systems, combining embedded optical sensor and additional light sources should be a great solution for portable device. Additional light sources can increase the accuracy of 3D positions detection. Embedded optical sensor based can provide a portable and continuous 3D interactive interface. Some 3D interactive systems based on this structure were designed and will be introduced in next chapters.

Chapter 3

T Mark Algorithm for Single User

In order to apply for portable devices, a system combined embedded optical sensor with additional light sources is chose as our platform. However, Disparity algorithm mentioned in chapter 2 only can detect 3D positions of single user in vertical direction. For friendly using, T Mark Algorithm was proposed to detect 3D positions (X, Y and Z)、rotation angle (θ) and tilt angle (φ). The principle、algorithm and experimental results will be described in this chapter.

3.1 Display with Embedded Optical Sensor

Embedded optical sensors can be integrated with the TFT LCD process. Display with embedded optical sensors can reduce system volume and eliminate the issues of blind range. Embedded optical sensor can receive the illumination of ambient light and transform into leakage currents. These currents are mapped to voltages which are divided into 10 bits data. However, the sensitivity of embedded optical sensor is still worse than camera. On the other hand, the resolution of embedded optical sensors is also limited to maintain the display image quality. For these reasons, the images captured by embedded optical sensors always are fuzzy and hard to calculate 3D positions of users.

3.2 T Mark System

Disparity algorithm [6] proposes a system combined additional light sources and embedded optical sensor to enhance the quality of images. This system is shown as Fig .3- 1.

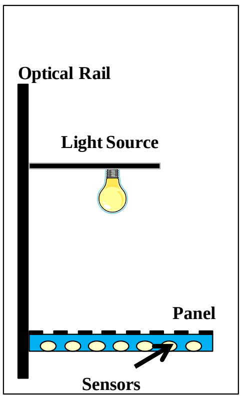


Fig .3- 1 Lighting Mode system [6]

However, Disparity Algorithm based on this structure only supports three degree of freedom (X, Y and Z). The tilt and rotation conditions cannot be detected. For widely using, we use light source with T shape pattern (as shown in Fig .3- 2) to interact with embedded optical sensor. By the way, five degrees of freedom (X, Y, Z, θ and φ) of light source can be obtained by analyzing the projection of light source. This system was called T Mark System and the T Mark Algorithm will be described.

Fig .3- 2 Light source with T shape pattern

The flow chart of T Mark Algorithm is shown as Fig .3- 3. Raw data of image will be captured by optical sensor. After de-background and de-noise, noise effects of ambient light and device noise can be eliminated. The clear image of light source will be obtained. If the tilt angle is larger than 20 degree, the projection of light source is very easy to be out of panel. This condition is not considered in our algorithm. Otherwise, we can use Full search, Windows search, Circle search and Ratio search to calculate 3D positions (X, Y and Z)、rotation angle (θ) and tilt angle (φ). Each search algorithm will be described clearly later.

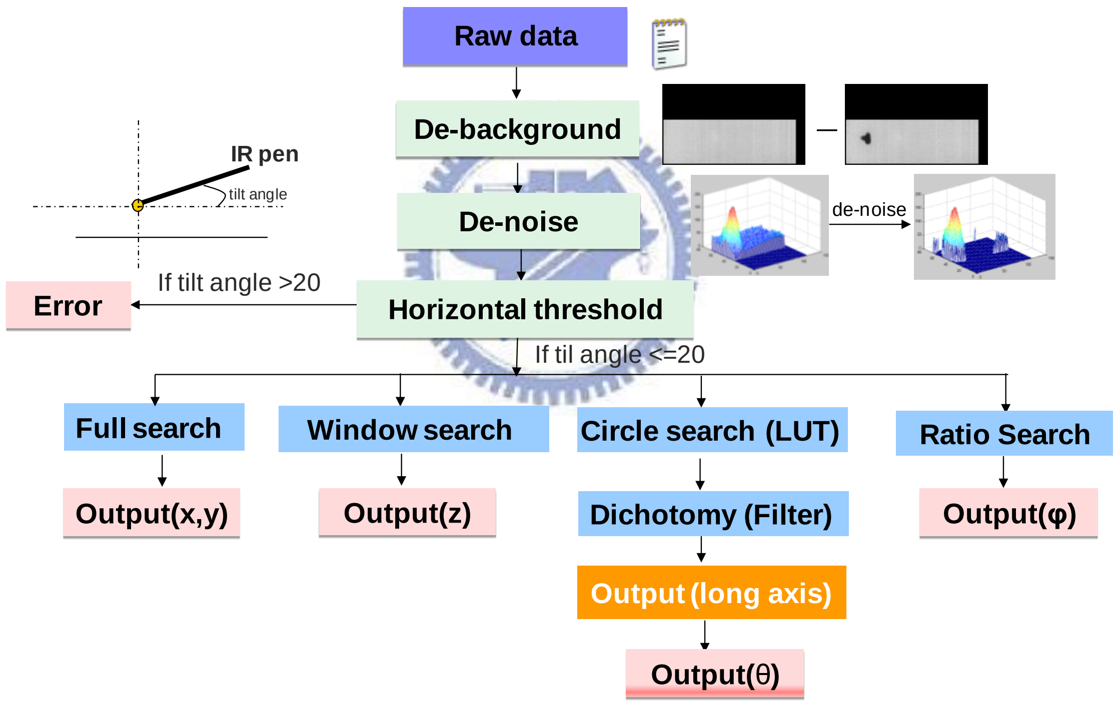


Fig .3- 3 T Mark algorithm

3.2.1 2D Coordinates (X and Y)

In general conditions, the maximum intensity of the projection always appears at the intersection of T shape (as shown in Fig .3- 4). In our algorithm, the position of captured image which has maximum intensity was regarded as the 2D coordinates (X and Y). However, strong noise of environment will affect the detection result of 2D coordinates. To avoid the effect of noise, Full search method was used to find the 2D positions (X and Y). Full search method uses a 3x3 matrix to calculate the sum of the values of each pixel and the adjacent eight points (as shown in Fig .3- 5). The position of maximum value in Full search will be regarded as the 2D positions. By the way, the effect of noise will be eliminated. The 2D positions can be determined accurately.

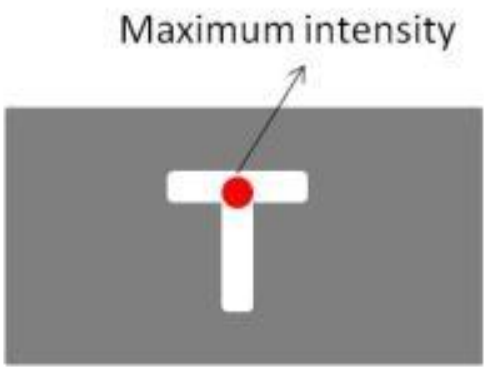


Fig .3- 4 2D coordinates of T shape

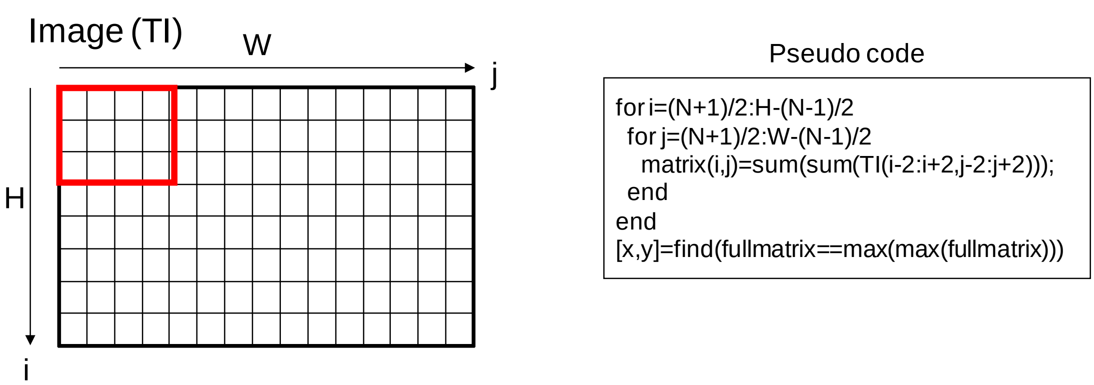


Fig .3- 5 Full search method

3.2.2 Height Information (Z)

Comparing with 2D touch system, 3D interactive system can distinguish more complex actions of user. However, how to detect accurate height information was a big issue if we just use bare finger to interact with system. With the help of light pen, height information can be calculated simply by the size of the projection of light pen. The area of projection is direct proportion to the height of light source within specific distance. In the other words, height of light source can be determined by checking the size of projection with look-up table. Window Search was designed to calculate the area of projection.

Window Search use different size of matrixes to cover on the 2D coordinates which were found by Full search and sum up the value of each pixel inside matrix. Next, the difference of adjacent matrixes can be obtained by using bigger one to subtract smaller one. For example, we use six different sizes of matrixes (A1 to A6) (as shown in Fig .3- 6) to detect the size of T shape projection. The difference A was obtained by using the sum within A2 matrix to subtract the sum within A1 matrix. The difference B was the difference between the sum within A3 matrix and the sum within A2 matrix (as shown in Fig .3- 7). Next, values of difference were normalized by dividing each difference by the number of pixel in each different range. If the window is bigger than T shape, the value of normalized difference should be very small. By the way, the size of T shape can be calculated.

Fig .3- 6 Different sizes of window matrixes

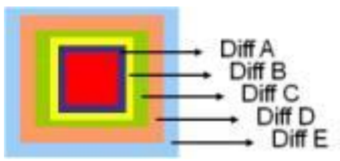


Fig .3- 7 Difference between matrixes

3.2.3 Rotation Angle (Θ)

Most of 3D interactive systems only detect three dimensions information (X, Y and Z). However, rotation was a common action in the interactive field. Although simple light source without pattern can be used to detect X, Y and Z information by Full search and Window search, but it cannot provide rotation information. For detecting rotation angle, pattern with direction is necessary. In our algorithm, the axis direction of the T shape can be used to find rotation angle. We established some angle filters (as shown in Fig .3- 8) to multiple by the T shape pattern. The maximum value after multiplying can help us to find the rotation angle.

Fig .3- 8 Angle filter for zero degree (left) and angle filter for forty degrees (right)

3.2.4 Tilt Angle (φ)

In addition to rotation angle, light source with T mark also can provide tilt angle (φ). T mark is consisted of two vertical axes. When the light source is not vertical to the panel, the ratio of two axes will change. The tilt angle can be determined after fitting the ratio with look-up table. The whole process called Ratio search method and the concept is shown as Fig .3- 9.

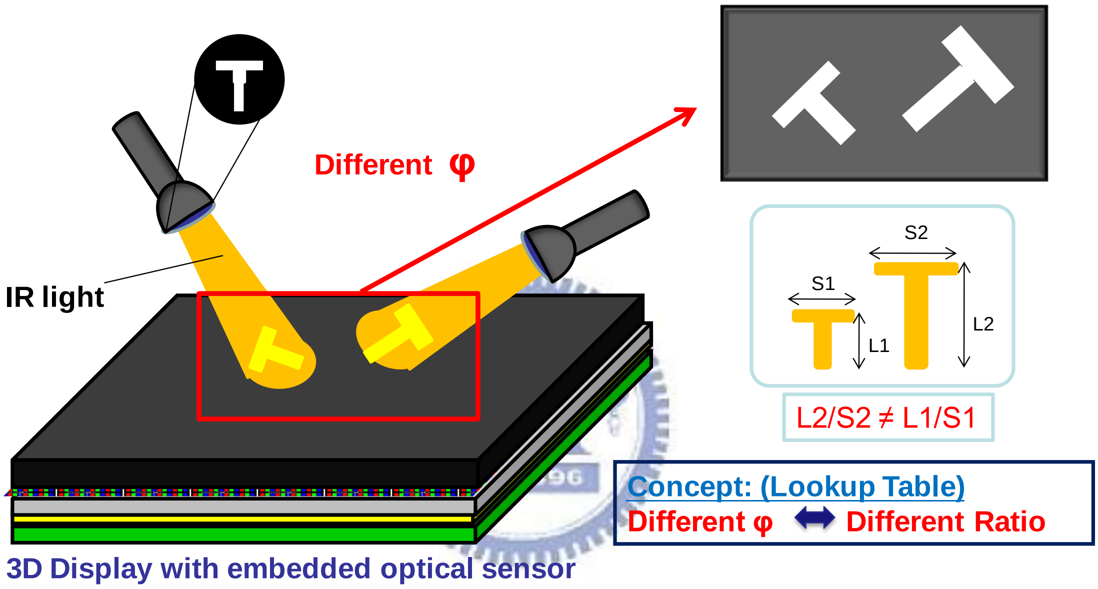


Fig .3- 9 Principle of Ratio search method

3.3 Experimental Platform

The experimental platform of T Mark System can be shown as Fig .3- 10. This platform consists of three major components: software develop platform (as shown in Fig .3- 11 (a))、3D display with embedded optical sensor (as shown in Fig .3- 11 (b)) and light pen with T mark (as shown in Fig .3- 11 (c)). The software platform is written by Boulder C++. This software platform can get image information from optical sensor and optimize some parameters of T Mark Algorithm in real time. The panel size of display is 2.83 inch. Due to the depth information of small size panel within 5cm, the working region in this system is designed from 0cm to 5cm.

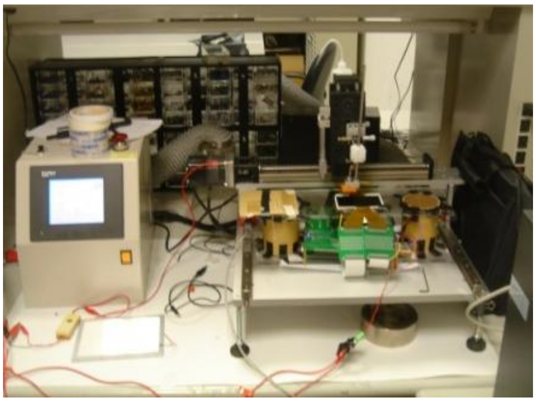


Fig .3- 10 Experimental platform

(a) (b) (c)

Fig .3- 11:(a) Software platform (b) Display with embedded optical sensor (c) T mark.

3.4 Experimental Results

Due to there were some line defects in our panel, some regions were cut off for better test result. The test results of X and Y coordinates were presented as Table .3- 1. The pixel size was 0.732mm. The error range only was a few pixels and the detection results were satisfactory. The detection results of Z information were also acceptable too. As shown in Table .3- 2, the accuracy of Z detection

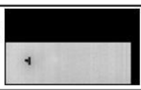
Unit: pixels (X↓ →Y)	x	y
Z=0cm 	48	19
Z=1cm 	46	18
Z=2cm 	44	18
Z=3cm 	44	18
Z=4cm 	44	18
Z=5cm 	45	16

Table .3- 1 Test results of 2D coordinates

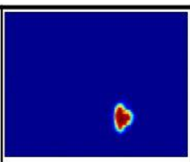
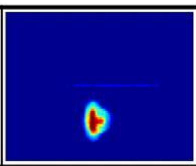
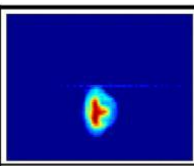
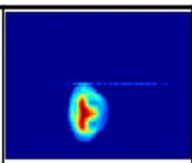
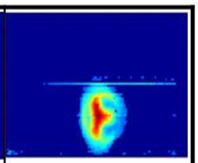
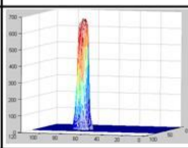
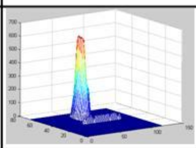
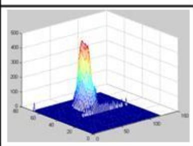
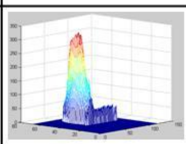
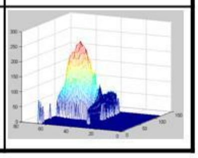
Panel 4-1 (Tilt angle=10°) De-noise=100 threshold=25 8.6V 0.36A					
Real	1cm	2cm	3cm	4cm	5cm
Result	1cm	2cm	3cm	4cm	5cm
Intensity					

Table .3- 2 Test results of Z

The detection results of tilt angle were always correct when the height of light source is more than 2cm. However, when we operated near panel, the detection result of tilt angle was incorrect (as show in Fig .3- 12). It may be caused by the low density of sensor. For this issue, the resolution of sensor needs to be increased for obtaining accurate detection result.

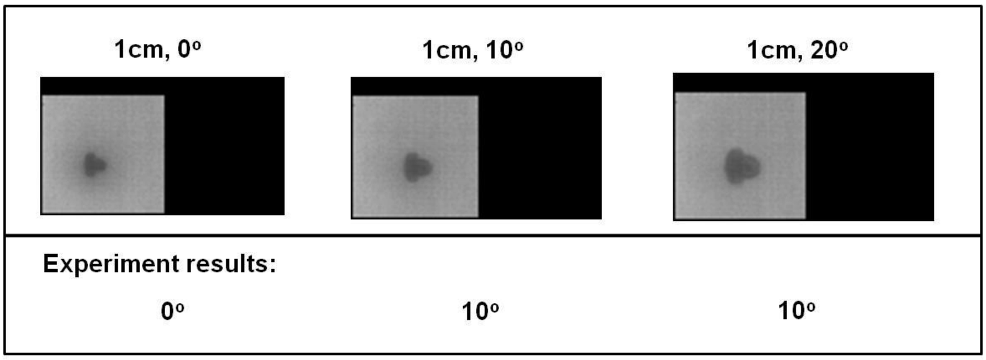


Fig .3- 12 Test results of tilt angle

3.5 Summary

T Mark System (as shown in Fig .3- 13) pasted T shape mark on light source to provide more information of light source. By the way, T Mark System can provide five degrees of freedom (X, Y, Z, θ and φ) of light source from 0cm to 5cm. Moreover, this system is suitable for using in be portable devices. However, this system still cannot support the multi-user or multi-touch.

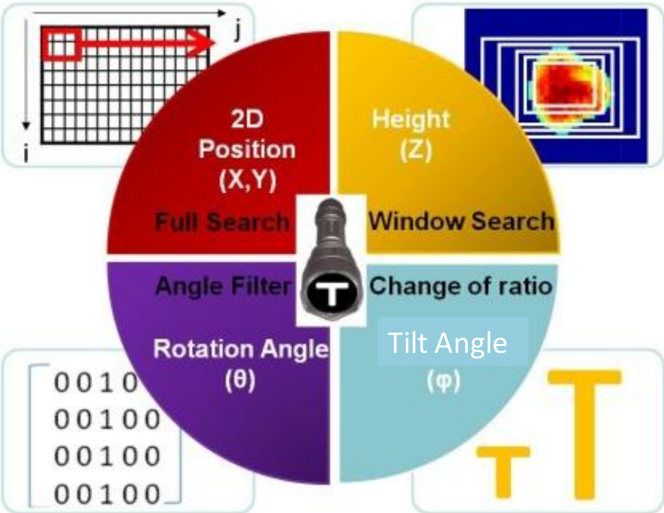


Fig .3- 13 Simplified diagram of T mark algorithm

Chapter 4

Color Filter Based for Multi-User

In the 3D interactive technologies which are designed on embedded optical sensor based, light source with T mark can provide five degree of freedoms (X, Y, Z, θ and φ) for 3D interaction. However, multi-touch and multi-user functions are very important for more complex interactive actions or using by multi-user at the same time. Color filter based will be proposed to provide multi-interaction and still can be using in portable devices.

4.1 Color Filter Based

For multi-interactive system, distinguish each user accurately is very important. However, if we just use several light sources to interact, the images captured by optical sensors will be very complex and the 3D positions were hard to be calculated. For this issue, image of each light source was needed to be distinguished separately. A simple and workable can be achieved by the color filters in LCD. Color filters (red、 green and blue) are band pass filter (as shown in Fig. 4- 1). The intensity of light with different wavelength will attenuate in different level when light pass through different color filter. By the way, if we place optical sensors under different kind of color filters, the intensity of light passing different color filter can be detected. On the other hand, red、green and blue LEDs were used as the interactive light sources. In this way, image of different light source can be obtained separately by optical sensors under different kind of color filters. The overall structure was presented as Fig. 4- 2.

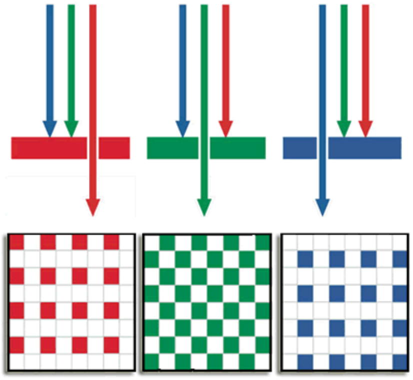


Fig. 4- 1 Color Filters are band pass filters

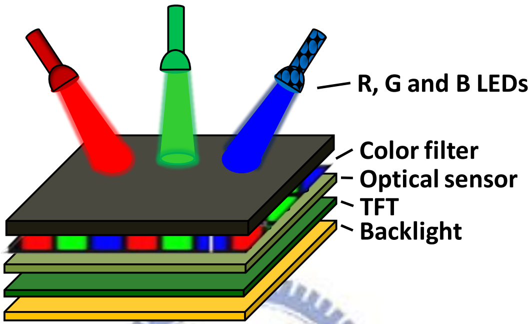


Fig. 4- 2 Structure of Color Filter Based

4.2 Characteristic of Common Color Filter

For checking the feasibility of color filter based, the characteristics of color filters need to be checked. Spectrum analyzer can be used to detect spectrums of color filters. At the beginning, white LED was tested to confirm the characteristic of each color filter. The spectrums of white LED and white light pass through each color filter were shown in Fig. 4- 3. We can find that different color filters can absorb different wavelength of light, but color filters cannot filter light perfectly in some waveband.

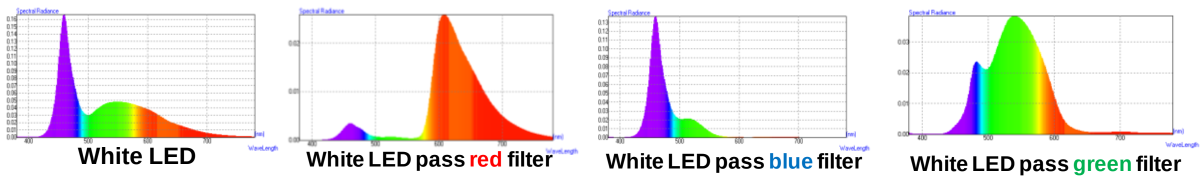


Fig. 4- 3 Spectrums of white LED and white light pass through each color filter

Next, the spectrums of common RGB LEDs passed through different color filter were presented in Fig. 4- 4. You can find that color filter still cannot filter perfectly because of the waveband of common LED is not narrow enough. If we use common LEDs to interact, the overlap conditions are inevitable.

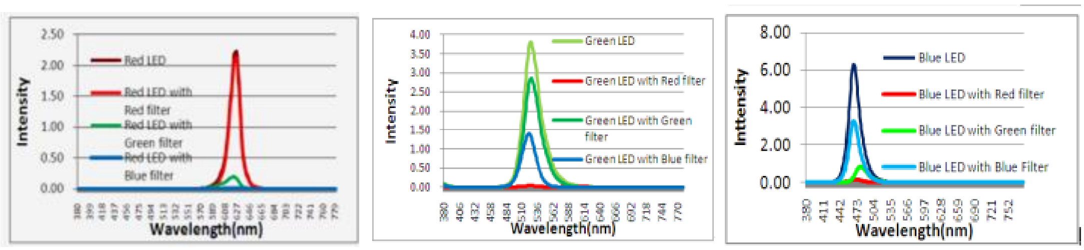


Fig. 4- 4 Different wavelength of light passes through different color filters

4.3 Algorithm of Color Filter Based

The 2D positions of each light source can be determined by Full search when the projections of each light source were separated by color filter perfectly. However, due to the imperfect match between common color filter and LEDs, the overlap conditions were inevitable. The more complex pattern we used, the more difficult to separate each pattern from overlap conditions. Finally, RGB LEDs without pattern were used to interact with Color Filter Based. Although this system cannot detect rotation angle (θ) and tilt angle (φ), the overlap conditions will be consisted of several circle and they can be distinguished easily by Hough transform.

Hough transform

The principle of the Hough transform is that perpendiculars to edge points of a circle cross in the center of the circle. Therefore, if we draw perpendicular lines to every edge point of our edge map, we should obtain bright `hot spots' in the centers of the circles. A circle with radius R and center (a, b) can be described as Eq. 4-1.

$$x = a + R\sin\theta$$

$$y = b + R\cos\theta \qquad \text{Eq. 4- 1}$$

For single circle with fixed radius, the search can be reduced to 2D. If we change the x and y coordinate into parameter coordinate (a, b), the location of (a, b) points in the parameter space can be expressed a circle of centered at (x, y) with radius R. The intersection of these circles is the true center of original circle (as shown in Fig. 4- 5). Multi-circles with fixed radius can be found with the same way. However, some incorrect circles maybe found (as shown in Fig. 4- 6). Those incorrect circles can be removed by comparing with the origin image.

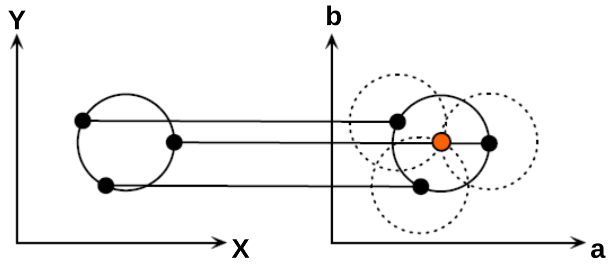


Fig. 4- 5 Circle with fixed radius

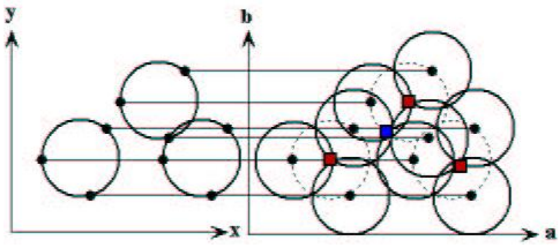


Fig. 4- 6 Multi-circles with fixed radius

In our system, we need to detect multi-circles with unknown radiuses. The critical concept is to establish a triple parameter coordinates (a, b, R). Each point (x, y) on the perimeter of a circle will produce a cone surface in parameter space (as shown in Fig. 4- 7). Each circle with unknown radius can be distinguished by using a three dimensional accumulation matrix.

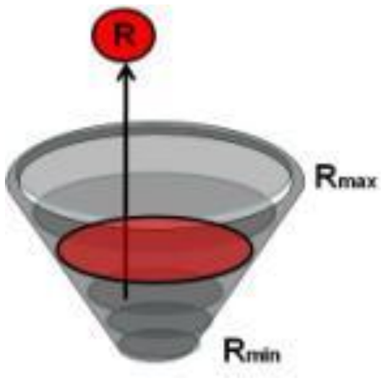


Fig. 4- 7 Multi-circles with unknown radius

Since the overlap issue had been solved, the Color Filter Based Algorithm will be proposed. The flow chart of algorithm is presented as Fig. 4- 8. In the real condition, optical sensors in our panel were embedded below red and blue color filters. For this reason, red and blue LED will be used in this algorithm. But this algorithm can be extended into three users by the same principle. At first, raw data of images are captured by red and blue sensor. Second, de-background can reduce the noise effects of ambient light and de-noise can further eliminate device noise. By the way, pure images of light source can be obtained. Next, Hough transform will be used to find possible center and radiuses of circles. After picking up some wrong circles, projections of light sources can be found. Finally, the projection of each light source can be distinguished by comparing the intensities captured by different optical sensors. The center of circle can be regarded as the 2D coordinates (X and Y)；Size of radius and value of maximum intensity at circle center can double check the result of height (Z). By the way, 3D positions of each light source can be obtained.

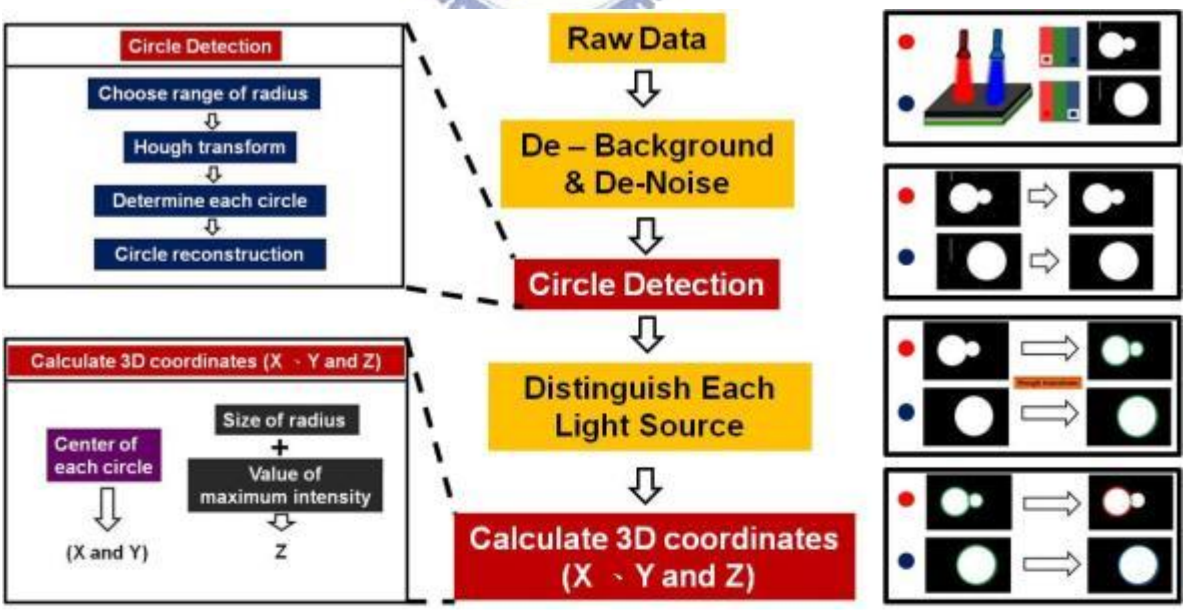


Fig. 4- 8 Flow chart of Color Filter Based Algorithm

4.4 Experimental Platform

The experimental platform can be shown as Fig. 4- 9. This platform consists of three major components: software development platform (as shown in Fig. 4- 10 (a)) 、display with embedded optical sensor (as shown in Fig. 4- 10 (b)) and LEDs (as show in Fig. 4- 10 (c)). The software platform can be used to get the captured image from optical sensor and be used to test Color Filter Algorithm. The size of display is 4 inch. In this system, different wavelengths LEDs were used to be interactive light source.

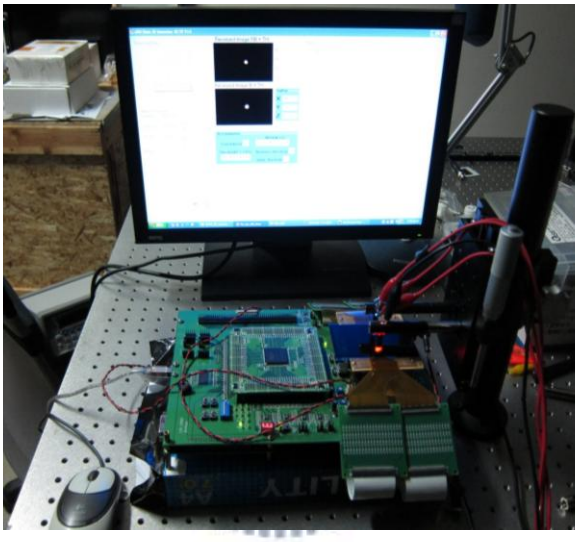
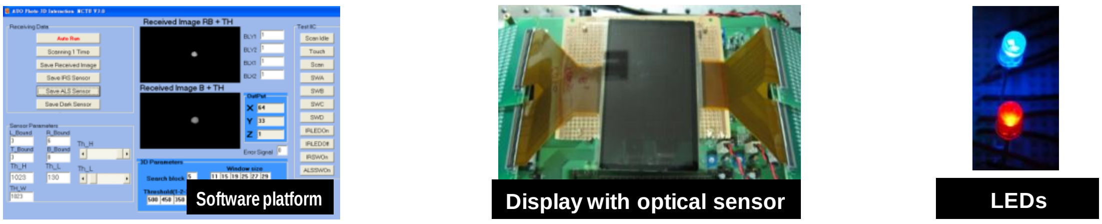


Fig. 4- 9 Experimental platform of Color Filter Based



(a) (b) (c)

Fig. 4- 10 (a) Software development platform (b) Display with optical sensor (c) Blue and Red LEDs

4.5 Experimental Results

At first, we needed to find the appropriate light sources for interaction. The sensing ranges of each light source on red sensor and blue sensor are shown in Table.4- 1. For larger working range and easier judgment of each light source, we choose red light and blue light to be interactive light sources. Next, the maximum intensities and size of radius of each light source at different height had been measured. The result at different heights (0cm to 5cm) under red and blue sensors is shown as Fig. 4- 11 and the size change of radius is shown as Fig. 4- 12. These cues can help us calculate the height of light source. In other way, red and blue sensors have different sensitivity to different wavelength of light. For better recognition, images captured by red sensors were be used to calculate the position of red LED and the images captured by blue sensors were used to calculate the position of blue LED.

Sensor \ Color of LED (Volt, Current)	Red (2V,0.05A)	Green (3V,0.01A)	Blue (3V,0.01A)	White (3V,0.01A)
Working range on IR sensor (Between light source and panel)	0~5cm	0~2cm	0~5cm	0~5cm
Working range on Blue sensor (Between light source and panel)	0cm	Over 5cm	Over 5cm	Over 5cm

Table.4- 1 Sensing range of both sensors with different LED

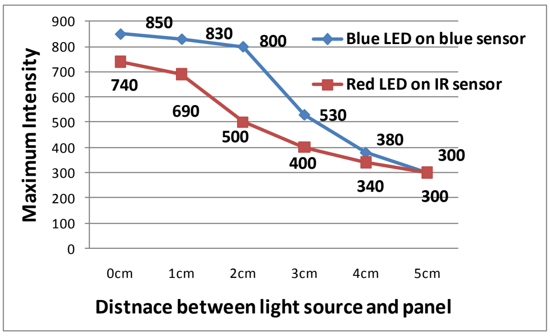


Fig. 4- 11 Maximum intensity at different height

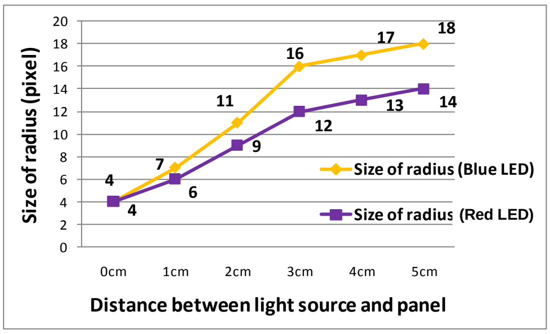
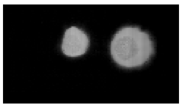
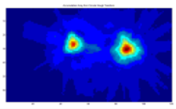
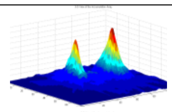
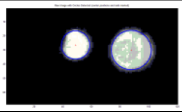
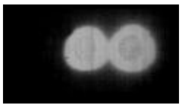
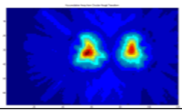
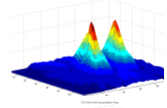
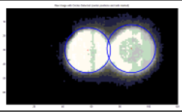
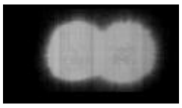
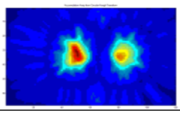
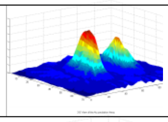
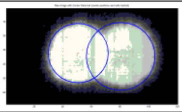
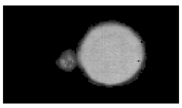
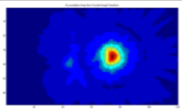
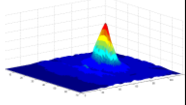
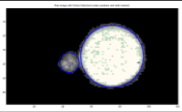


Fig. 4- 12 Size of radius at different height

Next, the accuracy of circle detection algorithm had been tested. The test results of four different cases were shown as Table.4- 2. These cases were circles with same size but with different degree of overlap and circles with different size. This algorithm is workable in different conditions. However, test result may be not correct when light is not vertical to panel.

	Scanned Image	**Accumulation**	**Accumulation Mesh**	**Detected Result**	**Center (X,Y) and Radius(pixel)**
(a)					**C1: (49,88) R1=10** **C2: (26,30) R2=14**
(b)					**C1: (30,57) R1=16** **C2: (29,88) R2=17**
(c)					**C1: (32,51) R1=21** **C2: (35,80) R2=24**
(d)					**C1: (39,34) R1=7** **C2: (45,74) R2=22**

Four cases: (a) Separate (b) Touch(c) Overlap (d) Circle with different size

Table.4- 2 Test result of circle detecting algorithm

Following was the experimental process of a random test (as shown in Fig. 4- 13). At first, two images captured by IR and blue sensor were obtained. Next, pure images of light source were obtained by de-background and de-noise. By circle detected, three circles (C1、C2 and C3) were found. Because of the distance between center of C1 and center of C3 was smaller than C1 and C2. C1 and C3 can be considered as the same light source captured by different optical sensors. After comparing the maximum intensity of C1 and C3, we can find that C1 was blue LED captured by by blue sensor, and C2 was red LED captured by IR sensor. Finally, 2D positions (X and Y) were identified by the center of each circle. Height information (Z) can be obtained by checking the radius of detected radius with look-up table. The detection of height can be confirmed again by the intensity at center of circle.

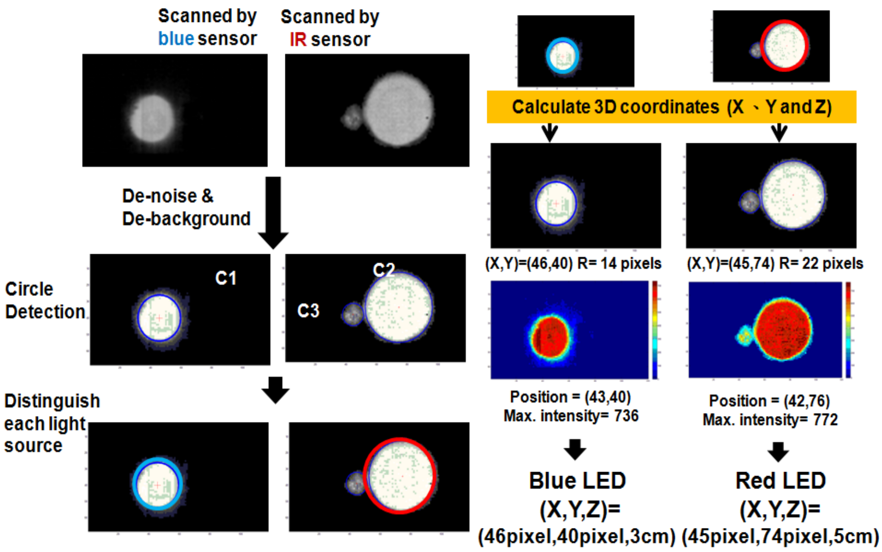


Fig. 4- 13 Experimental process

For verifying the accuracy of proposed algorithm, four different cases would be tested. Due to the working range of this system was 0cm to 5cm, these cases can standard for the extreme conditions. These results were shown in Table.4- 3. The detection results were good and the experimental results all fitted the real cases.

Cases	(1)Red in lower bound Blue in upper bound	(2)Red in upper bound Blue in lower bound	(3) Both LED at average height	(4) Overlap
Real condition (Z)	(1) Red LED : 0cm Blue LED : 5cm	(2) Red LED : 5cm Blue LED : 0cm	(3) Red LED : 3cm Blue LED : 3cm	(4) Red LED : 2cm Blue LED : 2cm
Image scanned by IR sensor				
Image scanned by blue sensor				
Distinguish each light source	IR sensor / Blue sensor	IR sensor / Blue sensor	IR sensor / Blue sensor	IR sensor / Blue sensor
Accumulation				
Circle detect				
Maximum intensity				
Center (X,Y) and Radius (pixels)	R : (30,39) r=4 B : (36,31) r=19	R : (41,38) r=14 B : (26,81) r=4	R : (50,81) r=12 B : (30,47) r=16	R : (32,102) r=7 B : (22,98) r=13
3D position (X,Y,Z)	R : (30,39,0cm) B : (36,31,5cm)	R : (41,38,5cm) B : (26,81,0cm)	R : (50,81,3cm) B : (30,47,3cm)	R : (32,102,2cm) B : (22,98,2cm)

Table.4- 3 Experimental results of four different cases

4.6 Summary

Color Filter Based not only provide the 3D positions of multi-touch or multi-user but also maintain original manufacturing costs. Using the characteristic of color filter, each light source can be distinguished. 2D positions (X and Y) can be calculated by the position of circle center. Height of light sources can be obtained by the size of radius and the value of maximum intensity. For the overlap issue, Hough transform was used to restructure each circle. By the way, a 3D multi-interactive system was achieved.

However, this system has some issues. For using the characteristic of color filter, the interactive light sources should be visible lights. The display image quality may be reduced when we watching and interacting with system at the same time. To overcome image quality issue, visible light may be replaced by IR light. However, Color Filter Based cannot distinguish each light source when using IR light to interact. Finally, another 3D multi-interactive system will be proposed in **Chapter 5**.

Chapter 5

Multi-Mark Based for Multi-User

In this chapter, multi-mark based will be described clearly. Multi-mark based was designed to solve the image quality issue of Color Filter Based by replacing visible light source with infrared light source. In order to distinguish these IR light source, different marks will be pasted on each light source to provide significant features. The 3D positions of each light source can be determined by analyzing the projections of these light sources. The structure、algorithm and experimental results will be describe in this chapter.

5.1 Multi-Mark Based

In Color Filter Based, color filters can help us to distinguish light sources with different wavelengths. However, Color Filter Algorithm will be fail when using IR light source to interact. In order to support multi-interaction and maintain the image quality, multi-mark based will be proposed.

The key components of multi-mark based are infrared light sources and several specific marks. Infrared light source can maintain the image quality of display when we watching and interacting at the same time. Light sources with different mark will produce different shapes of the projection. The 3D positions of each light source can be calculated by analyzing these projections. The schematic diagram of Multi-Mark Based was shown as Fig. 5- 1.

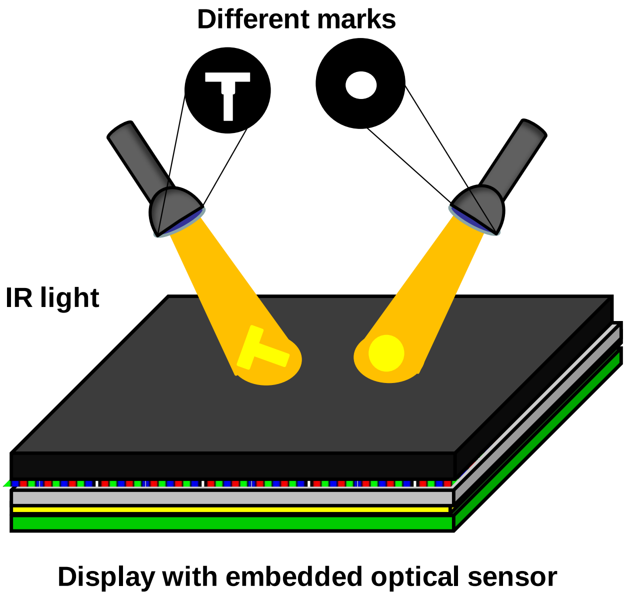


Fig. 5- 1 Schematic diagram of Multi-Mark Based

In order to determine each light source easily, the designed patterns should be significantly different to each other. Furthermore, it will be better if these light sources still can be distinguished when their projections were overlapping. Based on above requirements, three specific patterns were designed (as shown in Fig. 5- 2). These marks are consisted of line or circle. These types of image can be distinguished more easily. By combining each pattern with different infrared light source, Multi-Mark Algorithm which can support three users or three points was achieved. The complete algorithm will be presented in next section.

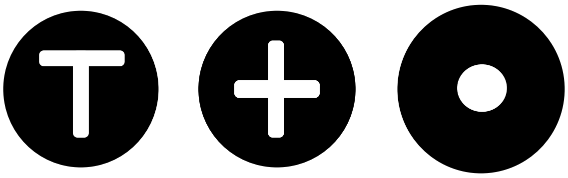


Fig. 5- 2 Selected Marks

5.2 The Algorithm of Multi-Mark Based

Multi-Mark Algorithm tries to distinguish each light source directly by some methods of image processing. By the way, this algorithm will be more complex than Color Filter Algorithm. The schematic diagram of Multi-Mark Algorithm is shown as Fig. 5- 3.

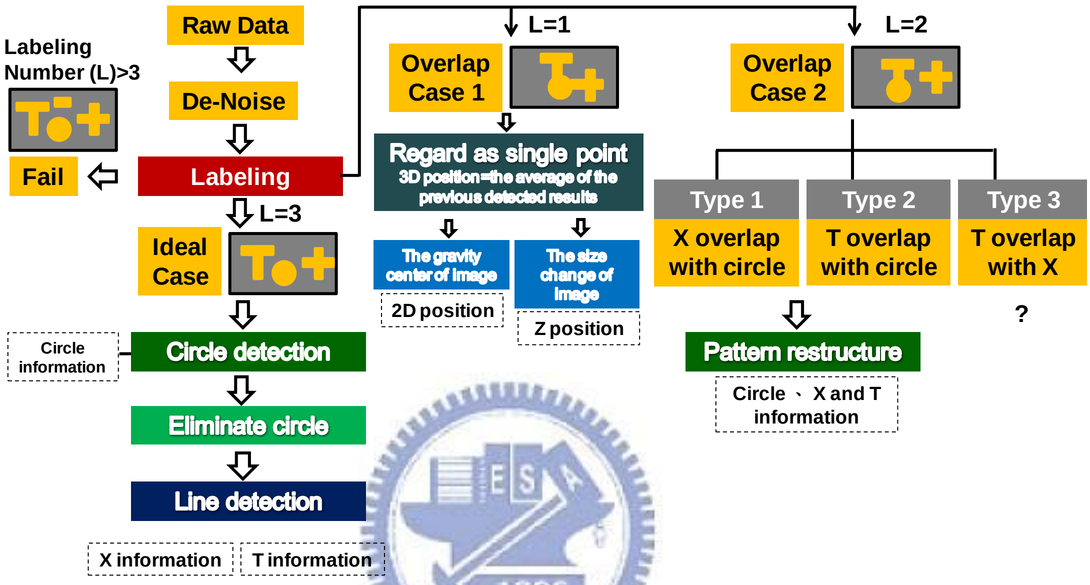


Fig. 5- 3 Multi-Mark Algorithm

At first, raw data of image will be captured by optical sensor. After de-noise, the device noise will be eliminated. Next, labeling will be executed to confirm the type of image. Labeling can label connected components. In the other words, labeling can tell us there are how many independent components in the image. An example is shown in Fig. 5- 4. There are two types of labeling: 4 connected components and 8 connected components (as shown in Fig. 5- 5). In our algorithm, 8 connected components will be chose for detecting some complex overlap conditions.

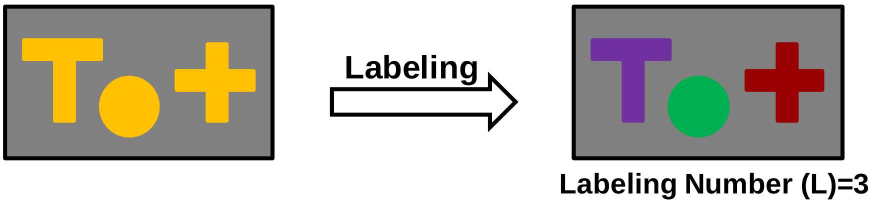


Fig. 5- 4 Example of labeling

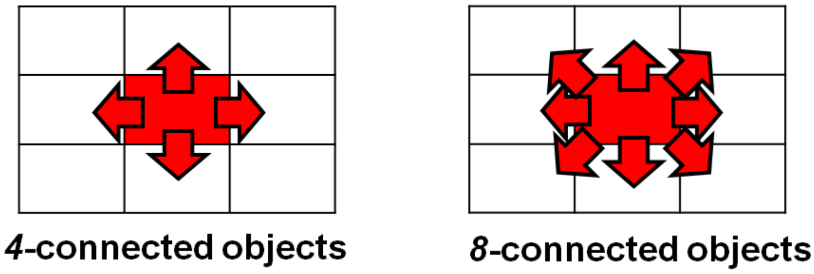


Fig. 5- 5 Two types of labeling

The condition of image can be confirmed after labeling. If the labeling number (L) is larger than three, it means there are some noises in captured image. This condition will be picked up form Multi-Mark Algorithm. If **labeling number is 3**, it means there are three independent components in captured image. It's also called ideal case. In this condition, circle detection will be executed to find the projection of circle mark. The process of circle detection can be expressed as Fig. 5- 6. Hough transform is used to find all possible circles in image. However, some unexpected circles will be found by Hough transform. These extra circles can be eliminated by judging the intensity at circle center. Finally, the projection of circle mark can be distinguished. The center and radius of projection also can be found.

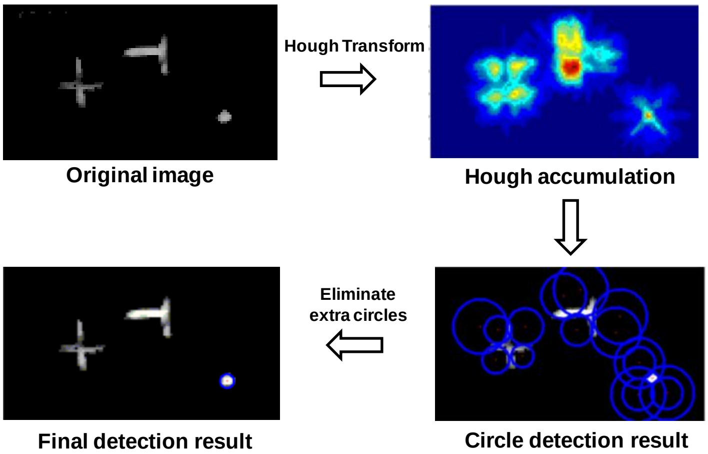


Fig. 5- 6 Process of circle detection

Circle detection can provide us the information of circle. The center of circle can be regarded as the 2D positions (X and Y). Checking the radius of circle with data, the height (Z) of light source can be determined.

Next, the projection of circle mark will be removed from original image to increase the accuracy of line detection. Line detection is used to find the projections of X and T mark. The flow chart of line detection is shown as Fig. 5- 7.

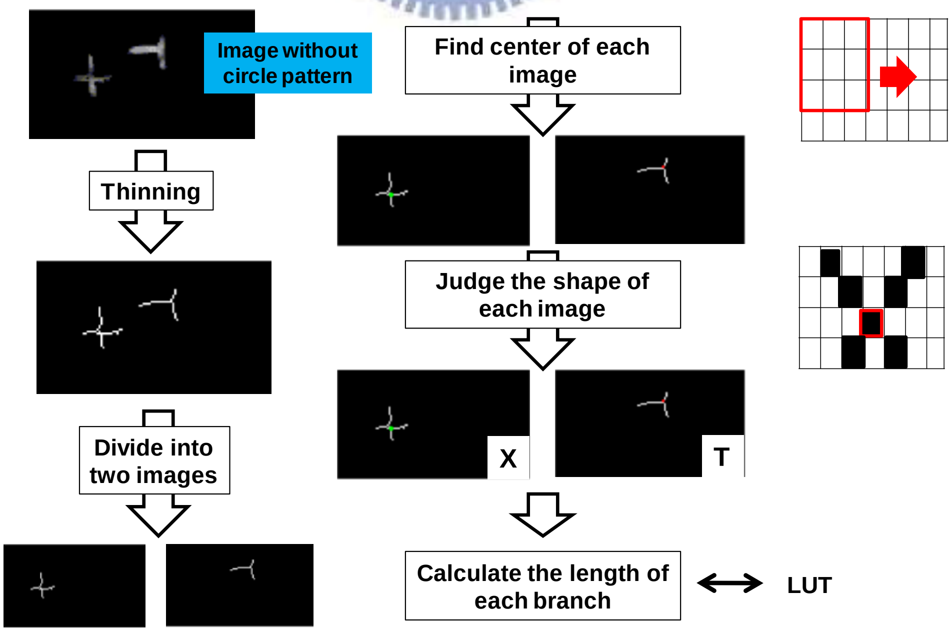


Fig. 5- 7 Flow chart of line detection

The projection of circle mark is eliminated form original image to avoid affecting the result of line detection. Next, thinning will be executed to get the line information of each pattern. As mentioned in T Mark System, Full search method can be used to find the center of image. However, Full search method is only workable for single light source. For this reason, each pattern will be saved as independent image. By the way, the 2D position of each pattern can be decided separately by Full search.

Since the 2D position of pattern was found, length of each branch of pattern will be accumulated from four different regions (as shown in Fig. 5- 8). Finally, the lengths of four branch of each pattern can be calculated. In general, the length of each branch in X pattern should be similar but obvious different in T pattern. These data can be used to calculate the heights of light sources pasted with X mark and T mark.

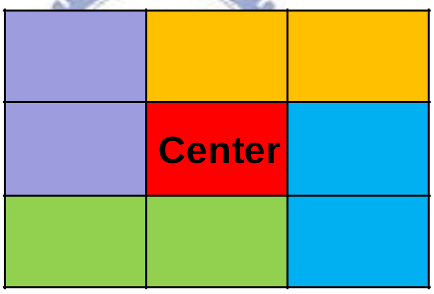


Fig. 5- 8 Four regions of branch searching

So far, the shape of each pattern still is unknown. There are two methods can help us to distinguish the shape of each pattern. At first, the distance between each point in pattern and the position of center will be calculated (as shown in Fig. 5- 9). After compiling statistics, the accumulation table of each pattern will be obtained. In general, X shape should have more points with the same distance than the T shape. On the other hands, we can use the length of each branch to double check the result of pattern recognition.

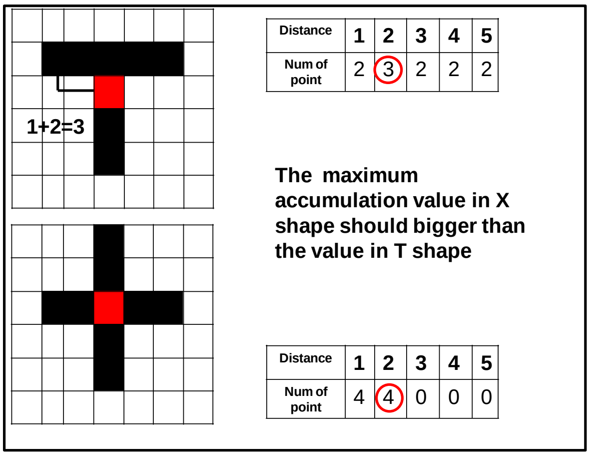


Fig. 5- 9 Principle of pattern recognition

Finally, the information of the projections of circle mark、X mark and T mark are obtained. The 3D positions of each light source pasted with different marks can also be confirmed in ideal case.

There exist other conditions after labeling the original image. If the **labeling number is one**, it means three projections of light sources are overlapping. In this condition, the 3D positions of each light source are very hard to be determined. This situation is always regarded as a single point in the field of 3D interaction. In Multi-Mark Algorithm, the condition is also regarded as a single point too. The gravity center of image will be regarded as 2D positions. Tracking the size change of whole image, the height information can be calculated.

If the **labeling number is two**, it should be partial overlap conditions. These conditions conclude three types of overlapping: (1) X shape overlaps with circle (2) T shape overlaps with circle (3) X shape overlaps with T shape. These conditions can be shown as Fig. 5- 10.

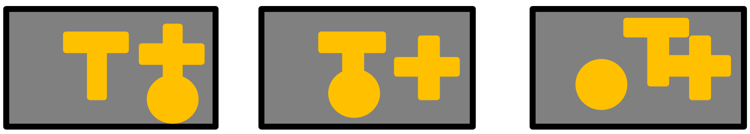


Fig. 5- 10 Three types of Overlap Case 2

At first, the projection of circle mark can be found by Hough transform. Next, circle pattern will be eliminated from the original image. Next, we will recognize the pattern which doesn't connect with circle. By the way, the type of overlap can be confirmed. If the detection result is type (3), it means T shape overlaps with X shape. Due to the characteristics of X shape and T shape are very similar, Multi-Mark Algorithm still has not find a perfect method to handle this condition.

Because of the maximum size of the projection of circle had been limited, the all overlap conditions of type (1) and type (2) can be demonstrated in Fig. 5- 11. These conditions can be distinguished after removing circle pattern and labeling again (as shown in Fig. 5- 12). Different overlap conditions will produce different new labeling number. Different methods are designed to restructure the pattern which overlaps with circle in different overlap conditions. These methods will be described clearly as following.

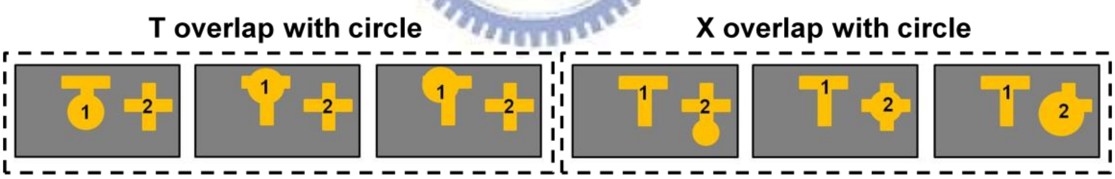


Fig. 5- 11 Possible conditions of type (1) and type (2)

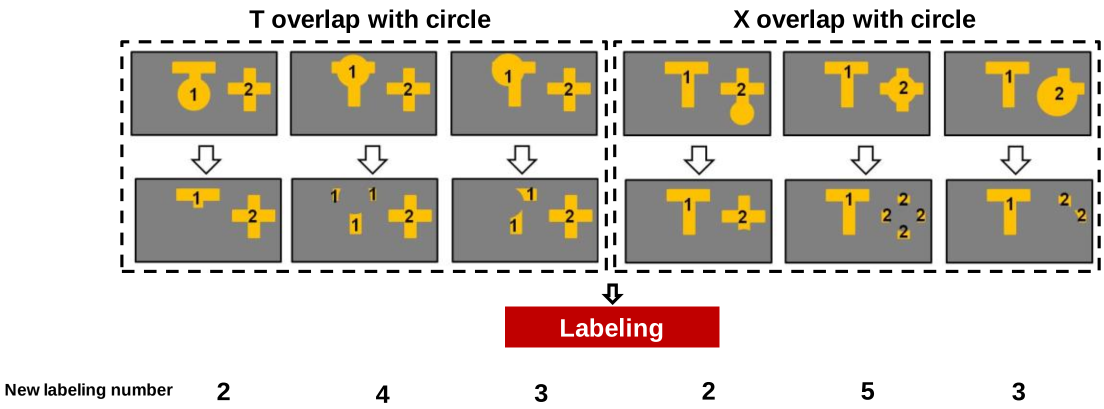


Fig. 5- 12 New labeling number of different kinds of overlap conditions

Restructure Method 1

This method is designed to restructure slight overlap conditions. Slight overlap conditions are expressed as Fig. 5- 13. In this condition, the pattern overlapping with circle still retains original center and at least one complete branch. The 2D positions of this pattern can be calculated by Full search which was mentioned in ideal case. The height can be determined by checking the length of complete branches with data.

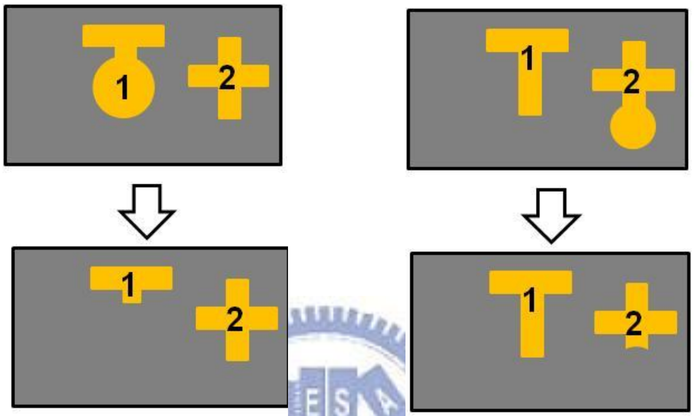


Fig. 5- 13 Slight overlap conditions

Restructure Method 2

This method is designed to restructure moderate overlap conditions (as shown in Fig. 5- 14). In this condition, the pattern overlapping with circle only keeps two partial branches. The center of pattern cannot be found directly. Hough line detection is used to find the most possible line of these remaining branches. The intersection of detected lines can be regarded as the center of pattern. The maximum distance between each branch and the center can be regarded as the length of each branch. The height can be determined by checking the length of each branch with data.

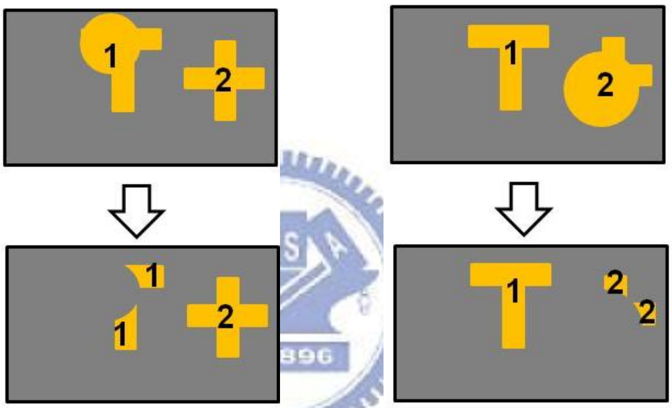


Fig. 5- 14 Moderate overlap conditions

Restructure Method 3

Sometimes, Hough line detection will get unexpected results when seriously overlap conditions (as shown in Fig. 5- 15). By the way, the restructure method 3 is proposed to handle these conditions. In this condition, the pattern overlapping with circle will remain part of every branch. The gravity of remaining parts will be used as the center of overlap pattern. The maximum distance between each branch and the center can be regarded as the length of each branch. The height also can be determined by data.

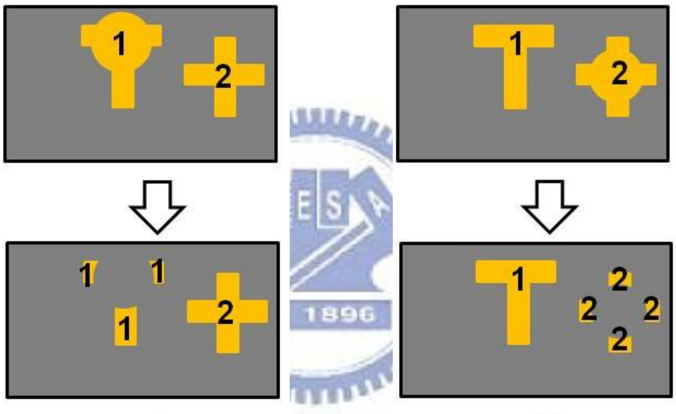


Fig. 5- 15 Seriously overlap conditions

In general, Multi-Mark Algorithm can distinguish the 3D positions of three light sources in most cases. However, sometimes Multi-Mark Algorithm will get unexpected results when the projection of T shape overlapping with the projection of X shape.

5.3 Experimental Platform

The experimental platform of Multi-Mark Based is shown in Fig. 5- 16. This platform consists of three major components: software development platform (as shown in Fig. 5- 17 (a)) 、display with embedded optical sensor (as shown in Fig. 5- 17 (b)) and IR light sources with different mark (as shown in Fig. 5- 17 (c)). The software development and display with embedded optical sensor are similar to the platform of Color Filter Based. Due to the size of panel is only 4 inch, the scattering angle of light source should be limited under 10 degrees to increase the work region.

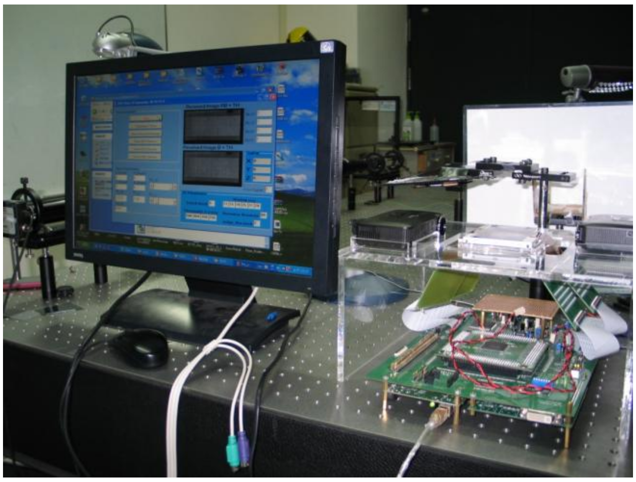


Fig. 5- 16 Experimental platform of Multi-Mark Based

(a) (b) (c)

Fig. 5- 17 : (a) Software development platform (b) Display with embedded optical sensor (c) IR light sources with different marks

5.4 Experimental Results

At first, the look-up table of each light source was established. The size of radius at different height of the projection of circle mark is shown as Fig. 5- 18. This curve is almost linear from 0cm to 5cm. The length of each branch at different height of the projection of T mark and X mark is shown as Fig. 5- 19 and Fig. 5- 20. These data were not linear when the heights higher than 3cm because of the intensities of these light sources are not strength enough. By the way, the working region of Multi-Mark Based was limited from 0cm to 3cm.

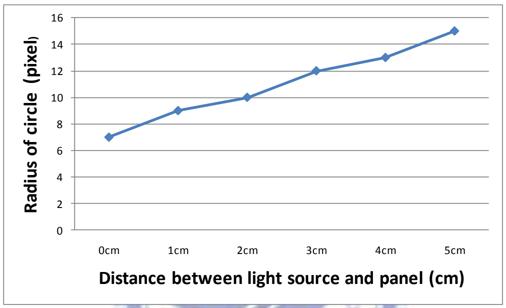


Fig. 5- 18 Size of radius at different height of circle mark

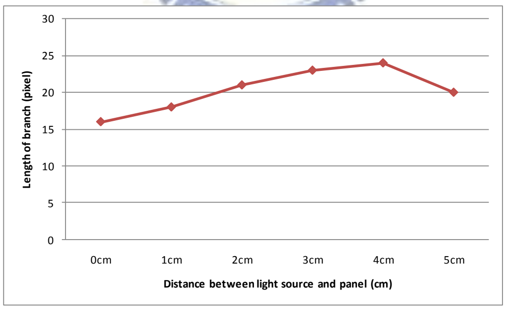


Fig. 5- 19 Length the projection at different height of T mark

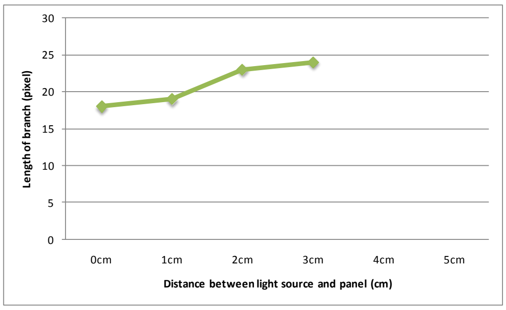


Fig. 5- 20 Length the projection at different height of X mark

Multi-Mark Algorithm divides the situations of captured images into three categories by the labeling number. Multi-Mark Algorithm was tested by the same context. Both random and extreme conditions were tested.

Ideal Case (Labeling number = 3)

Following is the experimental processing of a random test (as shown in Fig. 5- 21). At first, the captured image will be labeled to confirm the type of image. Next, all possible circles will be found by Hough transform. After eliminating extra circles, the projection of circle mark can be distinguished. By the way, the 3D positions of light source with circle mark can be determined by the position of center and the size of radius. Because of circle maybe affect the detection results of the others pattern, the image of circle will be removed. By the way, the remaining image will only compose with lines. Thinning can be used to enhance the characteristic of linear graph. Next, in order to avoid the impact of noise, Full search will be used to find the position of maximum intensity. After calculating the length of branches, the height information can be determined by checking the look-up table. By the way, the 3D positions of each light source can be obtained.

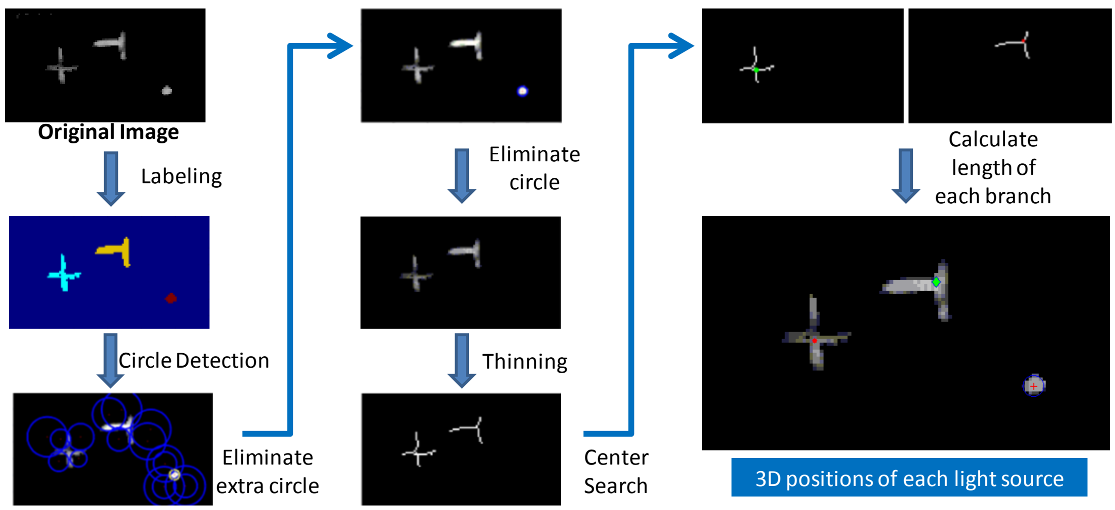


Fig. 5- 21 Experimental processing

In order to confirm the reliability of Multi-Mark Algorithm, three representative cases were tested (as shown in Fig. 5-22). These cases can be standard for the extreme conditions in our work region.

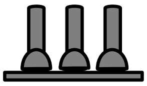
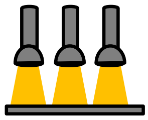
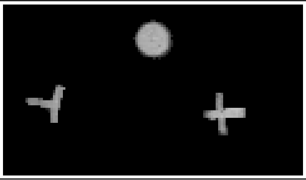
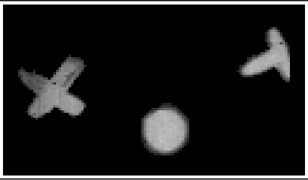
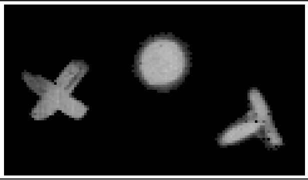
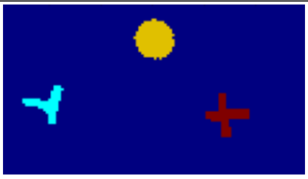
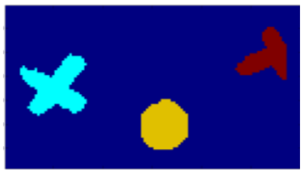
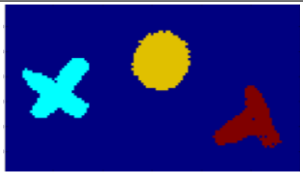
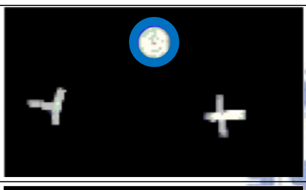
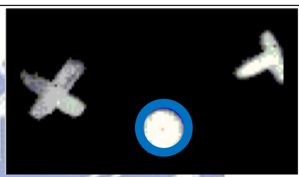
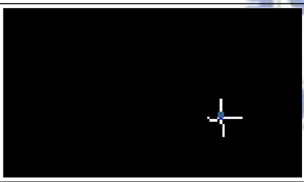
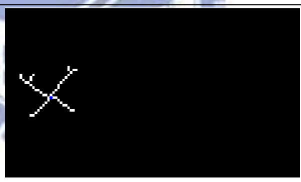
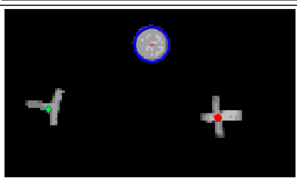
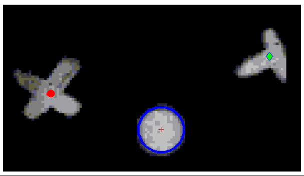
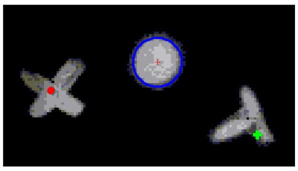
Z Real condition (cm)	 (O,T,X)=(0,0,0)	 (O,T,X)=(1,2,3)	 (O,T,X)=(3,3,3)
Captured Image			
Labeling			
Circle Detection			
X Detection			
T Detection			
 Final Result			
Detected 3D positions (pix,pix,cm)	X: (44,87,0cm) 0: (15,61,0cm) T: (41,19,0cm)	X: (36,20,3cm) 0: (50,65,1cm) T: (21,108,2cm)	X: (36,21,3cm) 0: (24,64,3cm) T: (54,105,3cm)
Real 3D positions	X: (42,87,0cm) 0: (15,61,0cm) T: (40,22,0cm)	X: (36,19,3cm) 0: (50,65,1cm) T: (21,108,2cm)	X: (35,23,3cm) 0: (24,64,3cm) T: (43,100,3cm)

Fig. 5- 22 Experimental results of three respective cases

The detection results fit the real condition substantially in the working region from 0cm to 5cm. The detection results of circle were accurate. However, the accuracy of 3D positions will reduce with the increase of the height of light source because of the image become more and more fuzzy.

Overlap Case 1 (Labeling number = 1)

In this condition, three light sources were regarded as a single light source. The gravity center of whole image will be used to standard for the 2D positions of whole image (as shown in Fig. 5- 23). The height information will increase 1cm when the area of captured image become 1.3 times of the previous image. This number was calculated from the experimental results.

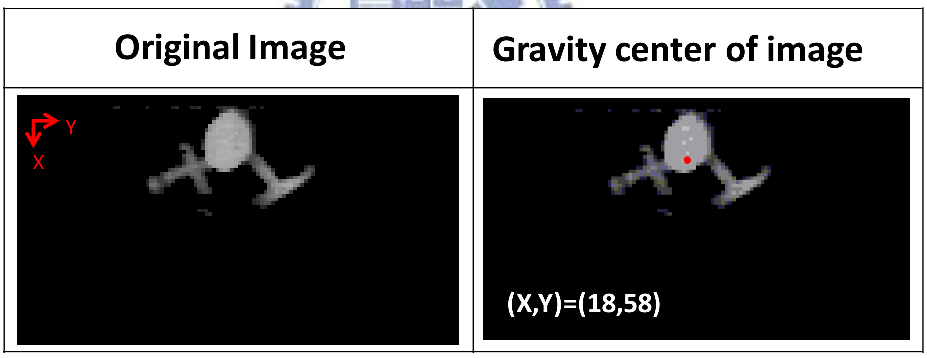


Fig. 5- 23 Experimental result of 2D positions

Overlap Case 2 (Labeling number = 2)

This case can be divided into three majors by the type of overlapping. Due to the characteristics of X and T are similar. Sometimes the detection results will be incorrect. On the other hands, the probabilities of seriously overlap conditions are very small when using in vertical direction. Two respective conditions are shown in Fig. 5- 24.

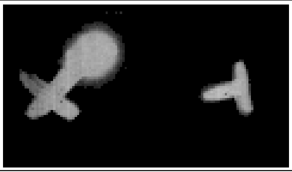
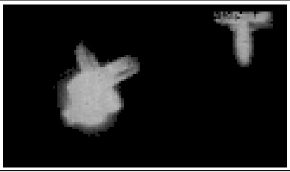
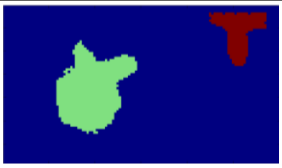
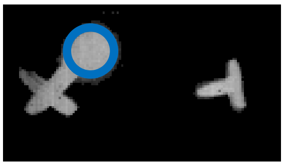
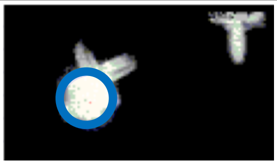
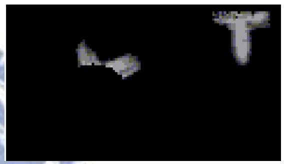
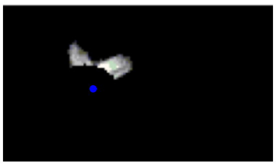
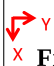
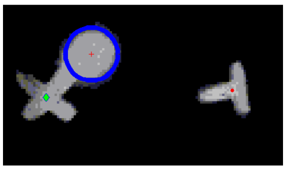
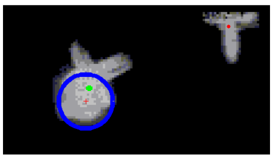
Z Real condition (cm)	(O,T,X)=(2,2,2) 	(O,T,X)=(2,2,2) 
Captured Image		
Labeling		
Circle Detection		
Eliminate circle		
Pattern (not connected with circle)		
Pattern (connected with circle)		
 Final Result		
Detected 3D positions (pix,pix,cm)	X: (39,20,2cm) 0: (21,39,2cm) T: (36,98,2cm)	X: (36,41,3cm) 0: (43,38,2cm) T: (10,101,2cm)
Real 3D positions	X: (39,20,2cm) 0: (21,39,2cm) T: (36,100,2cm)	X: (32,22,2cm) 0: (40,35,2cm) T: (8,105,2cm)

Fig. 5- 24 Experimental result of overlap case 2

5.5 Summary

Multi-Mark Based is a 3D interactive system which was proposed to solve the image quality issue of Color Filter Based. Different Marks pasted on each light source were used to distinguish each IR light source. Multi-Mark Based can detect the 3D positions of three users in most cases. However, the detection accuracy will reduce with the increase of the height of light source. To ensure the accuracy of this system, the working region is limited from 0cm to 3cm. This limit can be overcome by increasing the intensity of light source. On the other hand, Multi-Mark Based still cannot detect certain conditions perfectly. This algorithm will be optimized continuously.

Chapter 6

Conclusions and Future Works

6.1 Conclusions

3D interactive systems are potential interactive systems for next generation. Users can get convenient interactive interface and interact with 3D image more truly by 3D interactive systems. However, there are not suitable systems can be using in portable devices because of the requests of convenience and small volume.

In this research, three 3D interactive systems based on lighting mode system (combined with additional light sources and embedded optical sensor) were proposed. All of these systems can be applied on portable devices. T Mark System can detect the 3D positions (X, Y and Z)、rotation angle (θ) and tilt angle (φ) for single user from 0cm to 5cm. By the way, most of single interactive actions can be supported.

For supporting multi-interaction, two 3D multi-interactive systems called Color Filter Based and Multi-Mark Based were proposed. The comparison table of these two systems is shown in Fig. 6- 1. Both of these two systems can detect the 3D positions (X, Y and Z) for multi-touch or multi-user. Color Filter Based uses color filters to distinguish different visible light sources. This system is still workable when the projections overlap seriously. However, the visible light sources will reduce the display image quality when watching and interacting at the same time. Multi-Mark Based uses IR light sources pasted with different marks to replace visible light source. By the way, Multi-Mark Based can maintain the display image quality.

In general, both Color Filter Based and Multi-Mark Based can support for multi-user and multi-touch. However, based on their strengths and weaknesses, Color Filter Based is suitable for multi-touch system and Multi-Mark Based is suitable for Multi-user system.

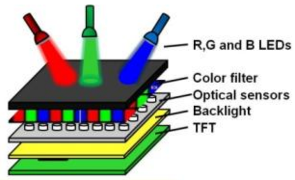
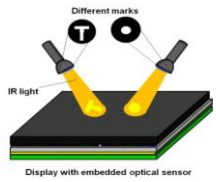
Color Filter Based	**Multi-Mark Based**
	
Support for multi-user and multi-touch	
Advantage: **1.Workable when seriously overlap** **Disadvantage:** **1.Reduce the image quality**	**Advantage:** **1.Maintain the display image quality** **Disadvantage:** **1. Overlap issues**
Multi-Touch System	**Multi-User System**

Fig. 6- 1 Comparison table of two proposed multi-interactive systems

6.2 Future Works

Until today, Multi-Mark Algorithm still suffers from some overlap conditions. Multi-Mark Based needs to be optimized for supporting common 3D interactive actions. The remaining overlap condition will be researched continually in the future. Moreover, the rotation and tilt angle will be detected in the optimized Multi-Mark Algorithm.

On the other hands, several light sources are different be taken for user. We will try uses light glove to replace light sources (as shown in Fig. 6- 2). By the way, a convenient interface can be established.

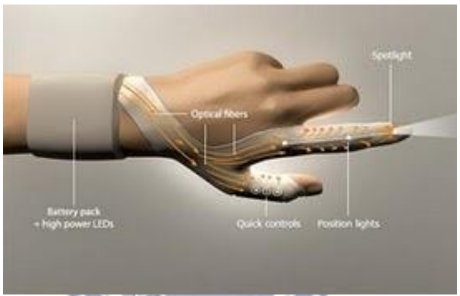


Fig. 6- 2 Light glove

References

[1] T Schlömer, B Poppinga, N Henze, ‘‘Gesture recognition with a Wii controller,’’ Proceedings of the 2nd international conference on Tangible and embedded interaction, 2008.

[2] JE Deutsch, M Borbely, J Filler, K Huhn, ‘‘Use of a Low-Cost, Commercially Available Gaming Console (Wii) for Rehabilitation of an Adolescent With Cerebral Palsy,’’ Physical Therapy, 2008.

[3] T Schou, ‘‘A Wii remote, a game engine, five sensor bars and a virtual reality theatre,’’ OZCHI, 2007.

[4] T Schlömer, B Poppinga, N Henze, ‘‘Gesture recognition with a Wii controller,’’ TEI 2008.

[5] J Giles, ‘Inside the race to hack the Kinect,’’ The New Scientist, 2010.

[6] JY Han, ‘‘Low-cost multi-touch sensing through frustrated total internal reflection,’’,UIST 2005.

[7] H Shimauchi, ‘‘Touch panel input device,’’ US Patent 4,812,833, 1989.

[8] Andrew Sears, and Ben Shneiderman, ‘‘High precision touchscreens: design strategies and comparisons with a mouse,’’ International Journal of Man-Machine Studies Volume 34, Issue 4, April 1991, Pages 593-613.

[9] M Zwicker, M Pauly, O Knoll, ‘‘Pointshop 3D: an interactive system for point-based surface editing,’’ SIGGRAPH 2002.

[10] T Mzoughi, SD Herring, JT Foley, MJ Morris, ‘‘WebTOP: A 3D interactive system for teaching and learning optics,’’ Computers & Education Volume 49, Issue 1, August 2007, Pages 110-129.

[11] Alex Olwal, Steven Feiner and Susanna Heyman, ”Rubbing and Tapping for Precise and Rapid Selection on Touch-Screen Displays,”CHI’08, pp295-304.

[12] D Wormell, ‘‘Advancements in 3D interactive devices for virtual environments,’’ Journal of Eurographics Workshop on Virtual Environments, 2003.

[13] R Lange, P Seitz, ‘‘Solid-state time-of flight range camera,’’ IEEE Journal of Quantum Electronics, 2001.

[14] Nicolai Marquardt, ‘‘The Continuous Interaction Space: Integrating Gestures Above a Surface with Direct Touch,’’ grouplab.cpsc.ucalgary.ca, 2009.

[15] Pai Tsung-Wei, ‘‘3D Interactive Display with Embedded Optical Sensors,’’ OPT 2008.

[16] N. Tarchanidis, N. Lygouras, ‘‘Data Glove With a Force Sensor,’’ IEEE Transactions on Instrumentation and Measurement, 2003.

[17] Virtual reality input devices, from the: http://sample.ctust.edu.tw

[18] JRW Morris, ‘‘Accelerometry--a technique for the measurement of human body

movements,'' Journal of Biomechanics, 1973.

[19] JC Lee, ''Hacking the Nintendo Wii Remote,'' IEEE Pervasive Computing, 2008.

[20] Low-Cost Multi-point Interactive Whiteboards Using the Wiimote, from the: http://johnnylee.net/projects/wii/

[21] S. J. Spencer, etal. ''A Low Cost Parallel Robot and Trajectory Optimization Method for Wrist and Forearm Rehabilitation using the Wii,'' BioRob, 2008.

[22] GJ Iddan, ''3D image in the stdio (and elsewhere),'' Proc. SPIE, 2001

[23] AD Wilson, ''TouchLight: an Imaging Touch Screen and Display for Gesture Based Interaction,'' ICMI, 2004.

[24] M Hirsch, D Lanman, H Holtzman, R Raskar, ''BiDi Screen: a thin, depth sensing LCD for 3D interaction using light fields,'' ACM SIGGRAPH Asia, 2009.

[25] M Ingebretsen, ''The Brave New Intelligent Interface,'' IEEE Intelligent Systems, 2010.

[26] W Yun-ping, ''Laser facula measurement technology using detector array,'' Laser Journal, 2007.

[27] Kauai, Hawaii, ''Photometric Stereo with General, Unknown Lighting,'' IEEE CVPR, 2001.

[28] W. Beil, I. C. Carlsen, ''Surface reconstruction from stereoscopy and "shape from shading" in SEM images,'' Machine Vision and Applications Volume 4, Number 4, 271-285.

[29] S. Burak Gokturk, Hakan Yalcin, Cyrus Bamji, "A Time-Of-Flight Depth Sensor - System Description, Issues and Solutions," CVPRW 2004 Volume 3.

[30] A Levin, R Fergus, F Durand, ''Image and Depth from a Conventional Camera with a coded aperture,'' ACM SIGGRAPH, 2007.