Multi-touch intelligent infrared interactive projection system

Wu Liming[1], Chen Hongna[1], Zeng Dezhi[1], Li Wenjie[1]

[1]Information Engineering School , Guangdong University of Technology，Guangdong Province, China
jkyjs@gdut.edu.cn,   31614792@qq.com   yyzengdezhi@163.com

Abstract - The system is based on optical principles of the multi-touch technology, multi-touch tracking, gesture recognition and other theories to design a laser pointer through the touch of a finger or a new type of light to operate a computer projection system to achieve low-cost, high-performance multi-touch Human-computer interaction technology, while taking advantage of augmented reality technology, and through computer network connections, to achieve "person - physical" information exchange

Index Terms - Multi-touch; infrared light; augmented reality; internet of things; gesture recognition.

1. Introduction

With the high-speed processing chip, multimedia and Internet Web technology rapid development and popularization, human-computer interactive research focuses on intelligent interactive, multi-model (multi-channel) - multimedia interactive, virtual interaction and the human-computer coordinated interaction etc people-centered huhuman-computer interaction.

In the interactive devices in recent years, there are many support new input device double-hand operation, which makes his hands and machine further application of natural interactions become possible. Such as the United States MERL (mitsubishi electric research laboratory) mond Touch Dia - [1], SONY Smart SkinE [2] lead, and Microsoft's latest launch of the Surface [3], etc. These interactive devices by identifying more touch, will be manually operated applied in human-computer interaction, and the results were very good. And interactions with the technology and development more and more [4], the common means has infrared DuiShe type, resistance membrane type, capacitive, acoustic type and the human-computer interaction based on camera such interactions.

Today's show way are mainly divided into display shows and projection display in two categories, display although is superior to the projection system to display, but with [5] every field in a display size of display devices and applications, take up the space for an increasing demand, display in some areas of higher advantage is gradually replaced projection display system [6]. But interactive projection system according to using the current technology, often have cost is too high, positioning prohibit, reaction speed is slow, planes, and disadvantages such as material limited only by infrared most product, the pen or electromagnetic pen to control, cannot achieve by hand can touch control, so not multipoint users to bring good experience. So people hope on the market can appear a can support more touching, also can achieve low cost, high precision, can interact in any plane projection system.

2. System Overview

This system is based on optical principle touch technology, the multi-point contacts tracking, gesture recognition theory, design the fingers or laser can through to operate the computer pen illuminate new projection system, the realization low-cost, high-performance multi-point touch huhuman-computer interaction. Also applied to Augmented Reality of augment Reality (AR) [7] the virtual Reality technology emerging field, and through the computer network connection, realize things -- the "that network information exchange function.

Specific means is, according to the actual application in vertical and horizontal need two plane take different scheme.
In vertical projection plane with a layer of manufacturing parallel infrared smooth, use camera to capture object when the planes that contact, again by infrared reflected the image analysis software to specific location, tracking, identification and interpretation through the projector, finally and express the feedback.
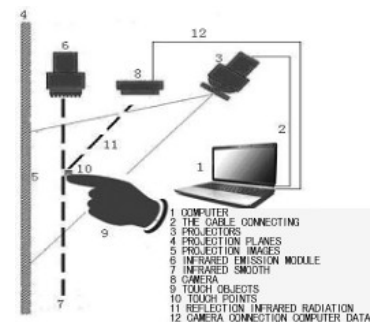

1 COMPUTER
2 THE CABLE CONNECTING
3 PROJECTORS
4 PROJECTION PLANES
5 PROJECTION IMAGES
6 INFRARED EMISSION MODULE
7 INFRARED SMOOTH
8 CAMERA
9 TOUCH OBJECTS
10 TOUCH POINTS
11 REFLECTION INFRARED RADIATION
12 CAMERA CONNECTION COMPUTER DATA

Fig 1 vertical planar structure

1 infrared leds
2 infrared diffuse light
3 acrylic board
4 contact objects
5 the DLP projection membrane
7 contact objects
8 the contact point
9 computer
10 computer connection of the camera data cables
11 cameras
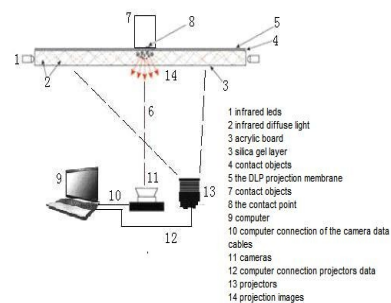12 computer connection projectors data
13 projectors
14 projection images

Fig 2 level graphic charts

Horizontal projection planes using transparent material. Projection planes of all round with infrared leds, infrared surrounded by closely in acrylic sheet diffuse within. Then according to the above vertical projection plan that complete the operation. Big difference is that in addition to identify the general ir window outside, still have the function of augmented reality. Cameras can identify placed on objects on planes, and information feedback about it to projection image.

3. The Hardware Equipments Buildings

The hardware equipments include to projection module, infrared emission module, sensing module and augmented reality module.

Projection module is a small profector connected to a computer. It puts the images, which originally shows in the display, in a general plane. The color of the plane cannot be too deep and its surface cannot be too concavo-convex, or it can't ensure the effect of profection images. And the horizontal plane should be made by transparent qualitative material. By repeated trials and compares, the colorless transparent acrylic board, of which thickness is 10mm, is the most cost-effecitve. Where is around the acrylic board should be polished so that infrared light can be shooted into the board. The plane for which touching of the acrylic board should be spraied a layer of smooth silica gel, and a piece of DLP projection membrane would be affixed above it to enhance the imaging result.

Infrared emission modeule is a row of infrared emission diode or linear shaped infrared laser. If using the linear shaped infrared lasser soulution, the general point shaped infrared lasser emitter, whose launch hole is affixed a linear shaped linear lens. Several laser emitter are fixed in a row, and between the adjacent is left certain intervals. If using infrared emission diode soulution, several infrared leds shouled be densely welded in a row, and a piece of paster is affixed to the upside and underside of it in order to prevent the infrared light, which may affect the projection images, from leaking as possible.

Sensing module is a high frame rate hd camera which can accrurately induct the infrared light reflected by objects touching the projection plane. Many cameras have filter lens which can filter infrared light. This filter lens is placed in the back of the camaera lens. It must be removed. Some infrared filter lens of the cameras can't be removed, then the camera lens should be replaced. Again a filter lens which can filter visible light is put in front of the camera lens. Then only infrared light can be into the camera. In addition, the camera selected should has good photosensitive effect in low light conditions. Among them, the horizontal plane is rear projected, the projector and camera are on the opposite side of the touching plane. And the vertical plane is front projected, the projector and camera are on the façade of the touching plane.

Augmented reality module is the specific two dimensional barcode printed in the bottom of some objects. These two dimensional barcode should be match to some corresponding object information and stored in the databases of the image analysis software[8]. In the actual operations, it should real-time response to the interact and display[9].

4. Image analysis and gesture recognition

The visible light of the image captured by special infrared camera has been filtered, what remain are some disturb points and some mass produced by the fingers touching the infrared plane.

Using image processing technology to process these video sequences. Firstly, extracting the motion pixels of the image by background subtraction method. Secondly, filtering the impurities of the background by binary method. Thirdly, filtering the disturb points by morphological filtering method and connecting motion pixels to become a connected area called mass. Fifthly, tracking these masses in the video sequences.

Gesture recognition means the corresponding instructions used to judge the movements of users' fingers, such as click, doubleclick, drag, amplification, narrowing, rotating, etc. The judge for fingers actions will normally come across in application program. Application program may be compiled by any programming language, such as C++, JAVA, or Flash. And the program used to process video sequences normally is compiled by C or C++ because of their high efficiency and good real-time implementation. In order to realizing the support of the program compiled by varieties of languages, the results of tracing touching points should be transfer to application program by certain standards. Now TUIO protocol has become one of these standards. In touching application program, a special module used to analysis the data transfer by TUIO protocol and translate it to corresponding instruction to transfer to other module is need.

The whole operation procedure of software part is as follows:

(1)Do image processing for the video sequences captured by camera and extracting the masses produced by fingers.

(2)Do motion tracking for the masses and record their trajctories..

(3)When a mass disappear, encapsulating its motion in-formation to the form of TUIO protocol and transferring it to application program.

(4)The TUIO analytic module of the application program analyses the information from the image processing module and transforms to corresponding instructions.

(5)Application program does response to the instructions

5. Image processing alogrithm

1) Background subtraction

Background subtraction method is one of the most commonly method which used to extracted the objects in video sequences. This method is particularly applicable for the conditions of fixed camera and the background color is relatively single. It's principle is: When we get a background frame, we use each pixel of the current frame to subtract the corresponding pixel in the background frame, if the result is 0,

we consider the pixel is a background pixel, if not , we consider that there is an object covered the background. There have many methods to constructing a background frame , in this paper we used the one order recursion filters to achieve it. We assume that the last frame is Fj-1 , last background frame is Bj-1 , then we can use the one order recursion filters algorithm to calculate the current background frame Bj like this: Bj=(1-a)Bj-1+aFj-1 ,the variable a is a self-adaption coefficient, and it's numeric area it from 0 to 1.

Assume that the background frame is $f_1(x,y)$, the current frame is $f_2(x,y)$,then, we can use the subtraction method to get the information of the motion pexils like this:

$$f(x,y) = f_2(x,y) - f_1(x,y) \qquad (1)$$

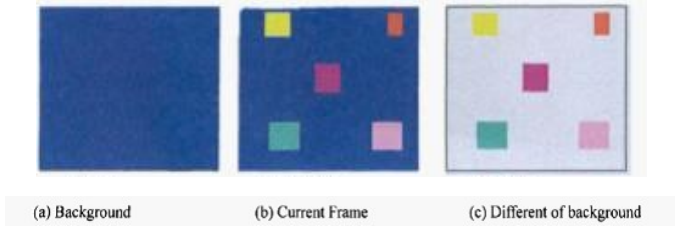(a) Background   (b) Current Frame   (c) Different of background

Fig 3   background subtraction

In reality condition , the background won't keep the same all the way , it will affected by the illuminance and impurity , if we just use the simple background subtraction , all the pixels will not be 0, so we should do binary processing after subtraction.

2) Binary processing algorithm on grayscale image

What we called Binary processing is to describe a image by 0 and 1. First ,we should set a threshold; Then we get the gray level of a pixel, if the level is smaller than the threshold , we set 0 for it , otherwise if the level is bigger than the threshold ,we set 1 for it .Assume that the background is $b(i,j)$, the current frame is $f(i,j)$, the motion object is $o(i,j)$, the threshold is th ,then the subtraction and binary processing can be described like this:

$$o(i,j) = \begin{cases} 1 & |f(i,j) - b(i,j)| > th \\ 0 & |f(i,j) - b(i,j)| < th \end{cases} \qquad (2)$$

The threshold th decided the tolerance of this algorithm to the interference factors.   The accuracy of getting motion objects is decided by whether the th is appropriate. When the th is too small , there will be a lot of noise points, which will be processed as motion blocks, and this will head to wrong touch event messages. When the th is too big, some motion blocks will not be detected , which will head to stupidity performance for the touch table. So user should be choose a appropriate threshold to make the binary processing reach to a best result.
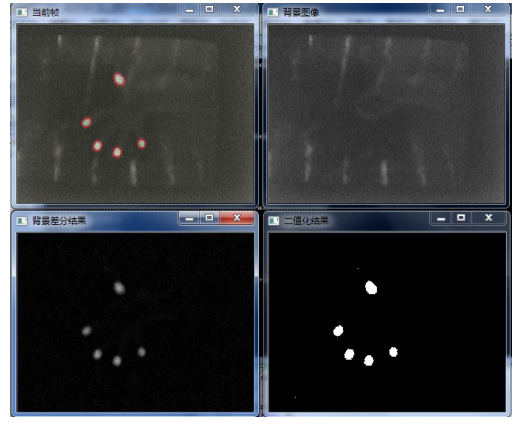

Figure 4: Result of threshold（th=30）

3) Morphology filter algorithm

Most effects of lighting will be removed by binary processing, but there still has some particle noises ,and there will be some holes in the motion blocks. By now, we can use the morphology filter algorithm to clear the noises.

The morphology filter algorithm contains two operations, one is called dilation, and the other is called erosion. These two kind of operations combined in different sequence can get two other operations, one is called open operation, and the other is called close operation. The open operation first do an erosion and then dilation, it is commonly used to clear noises. The close operation first do a dilation and then erosion, it is commonly used to fill the holes in object blocks.

The close operation is expressed by:

$$B \bullet S = (B \oplus S)\Theta S \qquad (3)$$

The open operation is expressed by

$$B \circ S = (B\Theta S) \oplus S \qquad (4)$$
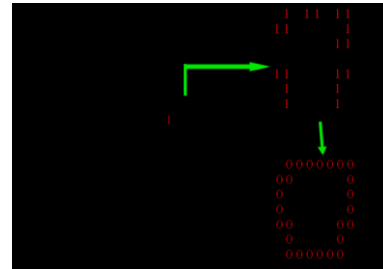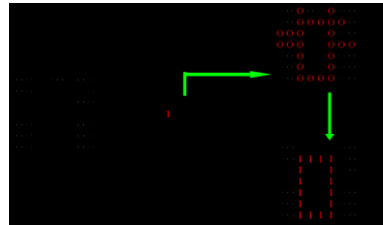

Fig 5   Close operation

Fig 6   Open operation

5*)* Object tracking algorithm

We can get some connected areas by the steps we discussed above, each connected area represent an object, the follow work is to track the objects in the frame stack, and record it's trajectory. As to describe an object, first we should do is to get the contour of the object. Such, one enclosing contour represent an object. The red lines in Figure 4 are the contours of motion objects.

For find the contours , we can use the canny algorithm. Assume that the two dimension Gaussian function

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

(5)

It's derivative of the second order on one orientation is:

$$G_n = \frac{\partial G}{\partial n} = \vec{n} \bullet \nabla G = (\cos\theta \quad \sin\theta) \bullet \begin{pmatrix} \frac{\partial G}{\partial x} \\ \frac{\partial G}{\partial x} \end{pmatrix}$$

(6)

Set: $\frac{\partial(G_n * f(x,y))}{\partial n} = 0$, $\frac{\partial(G_n * f(x,y))}{\partial \theta} = 0$ (7)

Then, we get the edge direction is:

$$\vec{n} = \frac{\nabla G * f(x,y)}{|\nabla G * f(x,y)|}$$

(8)

The edge intensity is:

$$A = |\nabla G * f(x,y)| = \sqrt{(\frac{\partial G}{\partial x} * f)^2 + (\frac{\partial G}{\partial y} * f)^2}$$

. (9)

When we get the contours, we can calculate a blocks centroid and area very convenience. We think the blocks too small or too big are noises or disturbances. We can set a range of area to filter the unreasonable motion blocks, to clear the disturbances again.

Because of the special application condition, we needn't to think about the shading and overlap, so the part of object tracking can adopt the simple centroid tracking. The procedure of this algorithm is:

1.Get the objects in the current frame, calculate centroid of each object;

2.Calculate the distances between each objects' centroid in the current frame to each objects' centroid in the last tracking queue, to get a matrix of distances;

3.Discriminant analysis. If a distance is smaller than a given value, we consider these two objects in different frame are the same object, then we add this centroid of object in the current frame to the object's trajectory. If the distances of an object in the current frame to any objects in last tracking queue are bigger than the given value, we consider this is a new motion object, and it represents the user pressed a new finger on the table, then we create a new tracking object and add it to the tracking queue. If the distances between a object in last tracking queue to the objects in current frame are bigger than the given value, we consider the object disappear in the current frame, this represents that a user's finger leave away from the table. For now, we can send the motion message of this object, and delete it from tracking queue.

6. conclusion

The emergence of multi-point touch, let people and system operation more easy and simple. It is not difficult to imagine the future computer interface must be abandoned keyboard, mouse these existing input devices, and direct in the image operate a computer. This work using low cost way to realize this goal.

In the actual building equipment test process, we found out that before the cameras, will filter cutoff wavelength reduced to just can let the red light, then we can be used by red laser pen for distance operation, solve the projection image area to touch hands as excessive. In addition, we also design a stylus have green red laser pens, two lamp holder with green laser to locate, reoccupy red laser to click and drag operation. Meanwhile, we also know that, as mentioned above level suitable for individual projection way office, leisure entertainment, and shall provide the miniaturization development; And the vertical projection way such as classroom is suitable for discussion and advertising, meeting, and shall such occasion to the large-scale development.

7. References

[1] Dietz P，Leigh D·Diamond touch：a multi—user touch technology[C]·Proceedings of the 14th Annual ACM Symposium on User Interface Software and Technology，Orlando，Florida，2001·

[2] Rekimoto J·Smart skin：an infrastructure for freehand I··pu—lation on interactive surfaces[C]·Proceedings of the SIGCHI conference on Human Factors in Computing Systems，ACM Press，Minneapolis，Minnesota，USA，2002·

[3] http://www.microsoft.com/surface.corn/index.html，2007，10·

[4] YuanBaoZong, RuanQiuQi WangYanJiang, etc, new generation fourth-generation human-computer interactive features the conceptual framework and key technologies, Chinese journal of electronics, 2003,31 (12A)

[5] Should root HuWenBo, QiuYong yulin,, etc, flat panel display technology. Beijing: people's posts and telecommunications publishing house, 2002:272-442.

[6] Eva Hornecker,"'I don'tunderstand it either,but it is cool'-Visitor Interactions with a Multi-Touch Table in a Museum",Horizontal Interactive Human Computer Systems,2008.TABLETOP 2008.3rd IEEE International Workshop on.2008 , 113 - 120

[7] Ronald T Azuma·A survey of augmented reality[J]·Presence：Teleoperators and Virtual Environments，1997，6(4)：355 - 385·

[8] Yingmin Zhang; Viard-Gaudin, C.; Liming Wu,An Online Hand-Drawn Electric Circuit Diagram Recongnition System Using Hidden Markov Models，Information Science and Engineering, 2008. ISISE '08. International Symposium on, 2008 , 143 - 148

[9] RanYang, ZhuFei, ChenKang, virtual reality and augmented reality technology industrial design application, wang, 2010,31,1 laser magazine

[10] Liming Wu; Qi Wang,A Recongfigurable Video Moving Target Deteciton IP Core,Digital Manufacturing and Automation (ICDMA), 2010 International Conference on，2010 , 104 - 108

[11] WuLiMing, CuiShanling, wang liping, LiuRunYu, repeat design chips new method, automatic detection precision engineering, optical 2008,16,5