A Transparent Display Interactive System with Artificial Intelligence Recognition

Jia-Chong Ho, Yu-Hsiang Tsai, Yu-Hsuan Su, Wei-Lin Hsu, Yu-Ying Lan, Sen-Yih Chou and Kuan-Ting Chen

jcho@itri.org.tw
[1] Electronic and Optoelectronic System Research Laboratories, Industrial Technology Research Institute, Chutung, Hsinchu, Taiwan 310, R.O.C.
Keywords: Image fusion, interactive system, transparent display.

ABSTRACT

In this research, we have developed a transparent display device with an information fusion interactive system. The interactive system has three main technologies: gaze tracking, object identification, and coordinate transformation for virtual image mapping. By integrating those technologies to achieve an AR-based interactive system on direct view transparent display.

1 INTRODUCTION

In recent years, the mixed reality interactive system is one of the main technology in the future. The users can manipulate virtual objects in real scenes, letting the users obtain the same real manipulation sense by augmented reality and sensor technology building. Actually, this technology is popular with the many field, such as medical, entertainment, education, etc.. We utilize AI (artificial intelligence) technology to implement object identification of user and object in the scene. Via face identification, the system will obtain the position of the user in the scene. Simultaneously, we used an AI training method to elevate the recognition rate for the object, it lets the system to recognize the object kind. Thus, an interactive operation based on the augmented reality with AI identification technology will be a develop topic in the future. In addition, transparent display technology is regarded as one of the main development technologies of the next generation display technology. Through sensing and positioning technology, virtual information can be displayed on the corresponding projection area of the screen. The information can be fused with background objects through a transparent display for augmented reality applications [1-2].

In this paper, we proposed an interactive system on transparent display, and we focus on how various functions can be properly integrated, along with its feasibility validation. The rest of the paper is organized as follows: the proposed dynamic interactive system, identification module, and transparent AMOLED display are described in section 2. Experimental results are reported in section 3. Finally, concluding remarks are given in section 4.

2 THE PROPOSED INTERACTIVE SYSTEM

In this section, we would introduce the proposed dynamic information fusion interactive system, including system structure, identification module, fusion image calculation, interactive interface design, and transparent display device. And, we apply this technology in two application field to demonstrate the feasibility of the proposed system. One is medical field, the other one is entertainment field. Fig. 1 shows a system flowchart of the application situation of the proposed interactive system. The user can directly watch the real object via the transparent display device, and the virtual digital image would show on the display.

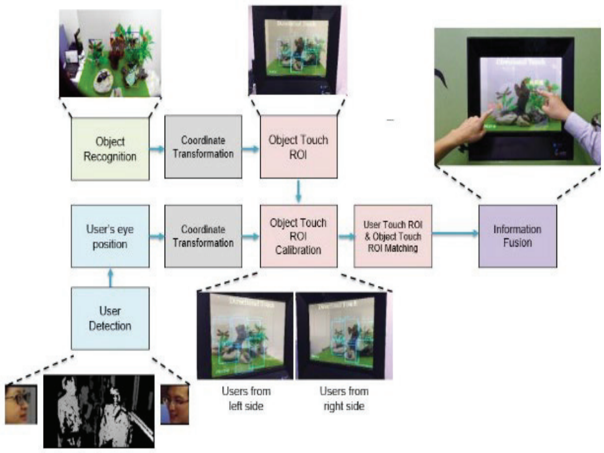


Fig. 1 The application situation of the proposed interactive system.

2.1 System Structure

Fig. 2 shows the software and the hardware structure of the proposed interactive system. The system peripheral has two cameras and transparent display. Via the in-camera device, the system would execute face identification and gaze tracking of the user. We used the out-camera device to identify the real object type and size. In the manipulation method design, we utilize an airbar to detect the touch position of the user, then shows the fusion image and data information results on the transparent display.

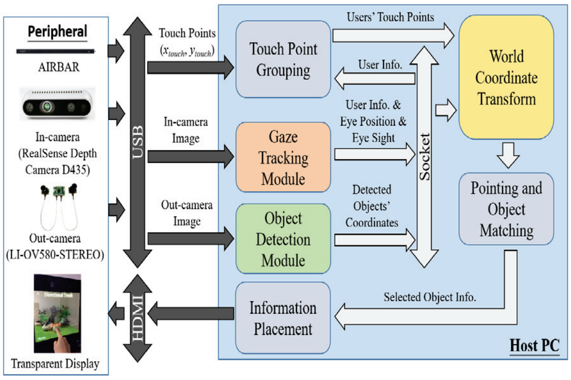


Fig. 2 The proposed interactive system structure.

2.2 Identification Module

In the identification module, we divided the system into two parts to discuss: one is user identification, the other is object identification.

(1) User identification

The gaze tracking module adopted the Intel RealSense D435 sensor to catch the image for face detection and facial landmarks. According to its high resolution on color image and stable measurement result on the depth, letting the gaze tracking module could predicted the facial landmarks of the user. To go a step further, the gaze tracking module could estimate the head pose, and could detect eyeball center and pupil center. Highly accuracy of head pose estimation, eyeball center and pupil center provide the gaze tracking module could calculate gaze vector. The gaze tracking module uses depth information from the depth sensor and adopted machine learning method to generate the identification results, and send the results to the host system by the socket communication automatically, its flowchart illustration as shown in figure 3. This module programming language uses Python, and it utilizes the Dlib library for machine learning. Beside this, in order to prevent user interference on system operation, the AI multiple feature recognition such as human pose recognition was integral in this interactive system also as shown in figure 4.

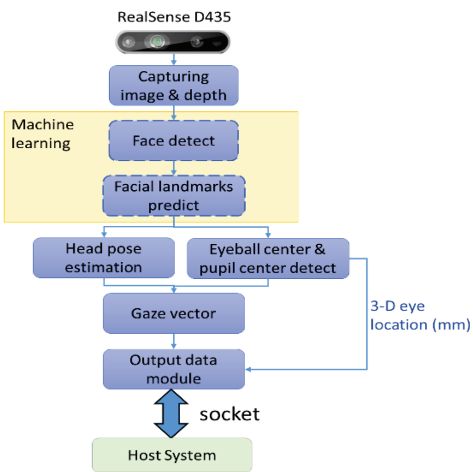
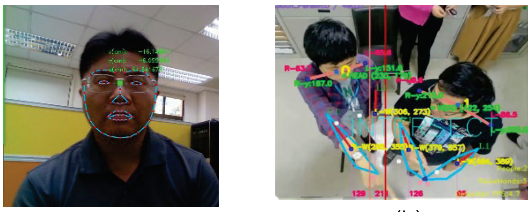


Fig. 3 The flow chart of user identification module.



(a) (b)

Fig. 4 The User's (a) gaze direction and (b) pose recognition was integral in system to identify user's behavior.

(2) Object identification

The object identification module adopted the stereo camera to catch the object image. We applied YOLO (You Only Look Once) [3-4] for the object identification module, which integrating object localization, feature extraction, and image classification. This algorithm uses end-to-end adopted the captured image and CNN (Convolutional neural network) [5] module to realize the object identification module. The method will generate several bounding-boxes and evaluate the probability of each bounding-box through a CNN module to identify the object and the position. Via the Socket communication, the host system would receive four parameters of the object image from the object identification module, those are central point coordinates (xo, yo) and its size (ho, wo). The figure 5 illustrates the flow chart of object identification.

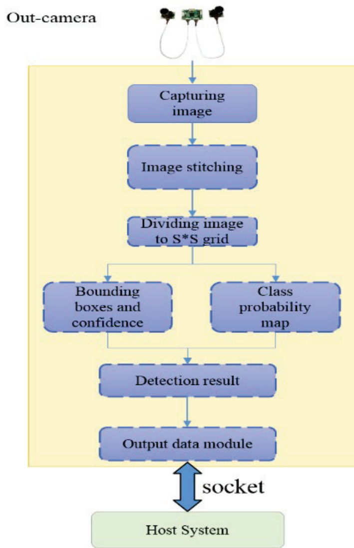


Fig. 5 The flow chart of object identification module.

2.3 Fusion Image Calculation

In the host system executes the data integration of the two modules, and show the fusion image on the transparent display device. However, the resolution of the camera is different to the transparent display. We used the formula (1) to calculate the deviation value of the image position.

$$\begin{bmatrix} S_x \\ S_y \end{bmatrix} = \begin{bmatrix} O_x * (S_{res\_w}/OC_{res\_w}) \\ O_y * (S_{res\_h}/OC_{res\_h}) \end{bmatrix} \quad (1)$$

where S_x and S_y is object position on the Screen, O_x and O_y is object position in the real scenes, and the transparent display and camera resolution are defined in (S_{res_w}, S_{res_h}) and (OC_{res_w}, OC_{res_h}), respectively. Via the gaze tracking function, the fusion image would follow the user position and gaze direction and change its angle and position. This application case is in the retail trade field, we adopted 32" transparent display to demonstrate the proposed interactive system, as shown in figure 6.

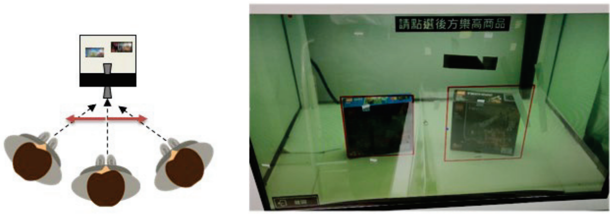


Fig. 6 The information fusion by user's gaze was test on 32" transparent showcase.

2.4 Transparent Display Device

In order to achieve vivid interaction effects, we adopted a high transparency 17 inch AMOLED (Active-matrix organic light-emitting diode, AMOLED) transparent display for the dynamic information fusion interactive system development. Fig. 7 shows the panel design schematic diagram of the transparent AMOLED display panel. Table 1 shows the detail specification of the AMOLED transparent display device.

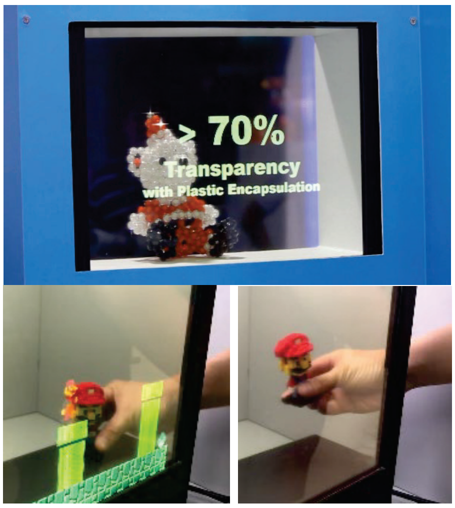


Fig. 7 The 17inch high transparency AMOLED.

Table 1. The Specification of 17 inch transparent AMOLED.

Item	Content	Unit
Panel size	17	inch
Physical dimension	4:3, landscape	
Resolution	47 PPI	
TFT type	LTPS-PTFT	
Transmittance of panel	70	%
Brightness	465	nit

3 Applications of Integrated System

This research has successfully developed a dynamic information fusion system and integral with a high transparency direct-view transparent display. Via the proposed technology implementations, the system would achieve the purpose of virtual information and real object fusion. This interactive system can be applied in a variety of application fields such as showcase window, retail cabinet, or medical education application, as shown in figure 8.

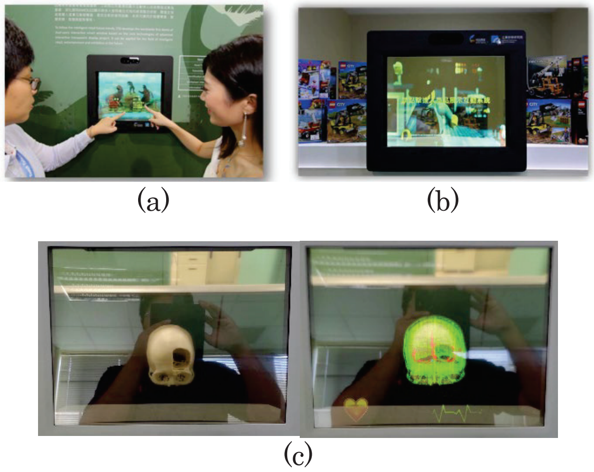


Fig. 8 The transparent display interactive system can be use on diversified application field such as (a) showcase window, (b) retail cabinet and (c) medical education.

4 CONCLUSIONS

In this research, we have developed a transparent AMOLED device with a dynamic information fusion interactive system based on AI recognition. This technology link the real scene objects and virtual information with a vivid experience. From the experimental results, the fusion image achieves a multi-view fusion image display by tracking the user's position and gaze direction. This technology can be used for several applications such as showcase window, retail cabinet, and medical education and bring a convenient future life.

REFERENCES

[1] Y. H. Lin, P. Y. Chen, K. K. Chen, K. Jen, H. H. Lu, C. H. Liu, and H. J. Hong, "Highly Transparent AMOLED for Augmented Reality Applications", Proc. SID Digest, 49, pp. 621-623 (2018).

[2] Y. H. Tsai, L. H. Wang, W. L. Hsu, and K. T. Chen, "Dynamic Information fusion system on Transparent Display", Proc. Flexible and Printed Electronics, 2019.

[3] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, Real-Time Object Detection”, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1-10, 2015.

[4] J. Redmon, and A. Farhadi, “YOLOv3: An Incremental Improvement”, in Proceedings of the *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1-6, 2018.

[5] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, Andrew Rabinovich, “Going Deeper with Convolutions”, in Proceedings of the *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1-12, 2015.