

# Development and Validation of Information Fusion Systems

Chieh-Han Yang<sup>1</sup>, Chien-Ju Li<sup>1</sup>, Yu-Hsiang Tsai<sup>1</sup>, and Mang, Ou-Yang<sup>2</sup>

<sup>1</sup>Electronic and Optoelectronic System Research Laboratories, Industrial Technology Research Institute, Hsinchu City, Taiwan, R.O.C

<sup>2</sup>Institute of Electrical and Control Engineering, National Yang Ming Chiao Tung University, Hsinchu City, Taiwan, R.O.C

## Abstract

Applications of transparent displays, such as the information fusion, have been implemented in various fields to replace conventional transparent glasses. For different scenarios, information fusion should be designed individually, requiring time to calibrate the components. To address this issue, a modular integration method is proposed. Key parameters are selected, and the design process is explained to construct the information fusion system. A simple experiment is conducted to validate the integration method, achieving an accuracy of 88 percent.

## Author Keywords

Argument reality, information fusion.

## 1. Introduction

Display technology is evolving rapidly, and the applications of transparent displays have become increasingly popular. The transparent display offers a unique opportunity to merge the digital and physical worlds, enabling applications such as augmented reality [1], smart windows, and interactive surfaces. These applications require careful consideration of the relationship between humans and objects, making them particularly relevant in the field of information fusion which is plotted in Figure 1. With significant advancements in computational hardware, real-time object and facial recognition are no longer technological bottlenecks. Information fusion techniques such as augmented reality, mixed reality [2], and smart windows, which combine transparent displays with AI, also have captured the interest of researchers.

On the other hand, in order to make the application more flexible and adaptable to various scenarios, the display has been designed to be rotatable and supports multiple viewing angles. Take two simple cases for example, landscape and portrait modes: Landscape mode is the typical usage for most fixed displays, where no rotation calibration is required for any parameters. On the other hand, portrait mode is rotating the landscape mode 90 degrees. Unlike the tablets or smartphones, the application on the transparent display cares more about the spatial position. Therefore, no scaling or calibration is needed after rotation and the actual position is adopted to show the information fusion content. Both landscape mode and portrait mode of a rotatable display are illustrated in Figure 2 and Figure 3.



Figure 1. The information fusion task on the transparent display.

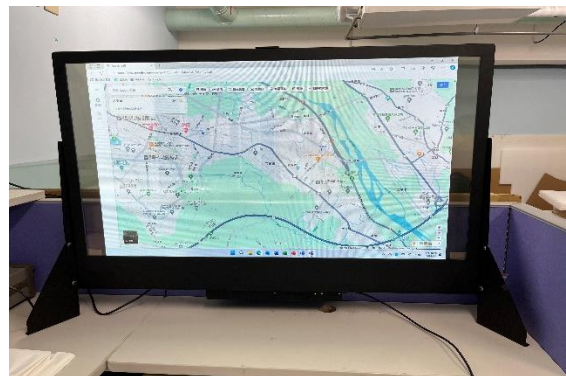


Figure 2. Landscape mode of Rotatable transparent display.



Figure 3. Portrait mode of Rotatable transparent display.

## 2. Modular parameters

In this section, the necessary parameters would be explained. Figure 4 plots the schematic diagram of an information fusion application. A basic information fusion application is composed of an observer, an object, and a transparent display. The key nodes of the human eyes are detected to represent the observer. The object could be detected by a camera or preset in a specification location relative to the display or even GPS information.

To modularize the information fusion task, the key technique is to transfer the axis of all elements to the same coordinate system. In this paper, the origin is set to the center of the upper edge of the display panel, and the integrated system follows the coordinate of right for positive x and up for positive y while facing to the display from the observer side. The relative distance from all the elements except the computing device to the origin should be set as parameters and measured in advance to adapt to different cases. If the axis direction of the camera does not fit the integrated system, the captured screen is suggested to fix in the order of rotating the result and then shifting the center to the desired location.

In tradition, the pose of the camera should be fixed and usually designed parallel to the display panel to build an easy-compute model. However, there will be some orientation difference during the setup procedure. In general, the small angel difference would be ignored if the object looks big enough in the transparent display, otherwise the system builder would adjust the camera manually to correct the information location on the display.

In this paper, the roll, pitch, yaw, the actual size, and the field of view (FOV) of the camera are set as initial parameters to calibrate the content location automatically. The pose of the camera can be captured by measurement physically, by sensor or by reading the data from its inertia measurement unit (IMU). For a single camera, rotating right is equivalent to spinning the world left. Therefore, inverting the data with the angle difference to the display and then shifting the origin to the origin of the integrated system finishes the calibration of a camera.

For a common flat-panel display, the parameters are the resolution, the size of the panel, the distance in the three axes of the top-left corner of the panel to the origin, and the placement degree of the display. When the display is placed in landscape mode or portrait mode, the only focus is where the pixel (0,0) is in the integrated system coordinate. The other display related data such as axis changes or the content location could be calculated by these parameters.

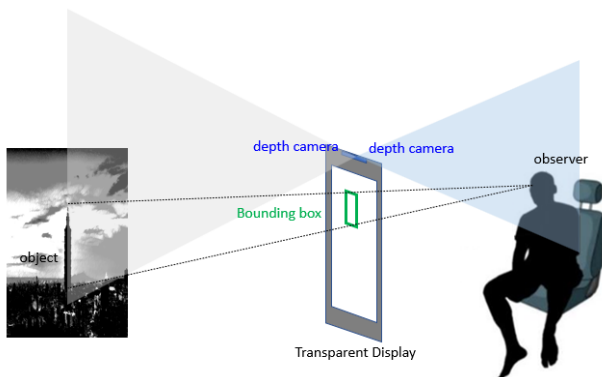


Figure 4. The schematic diagram of a basic information fusion application

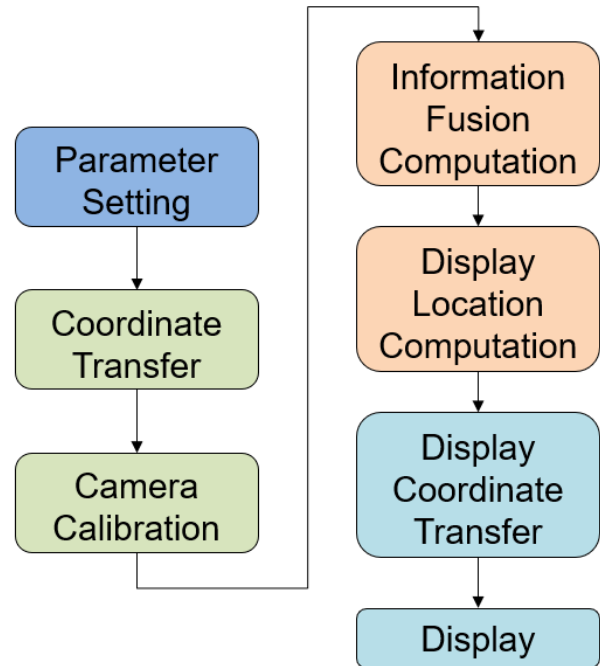
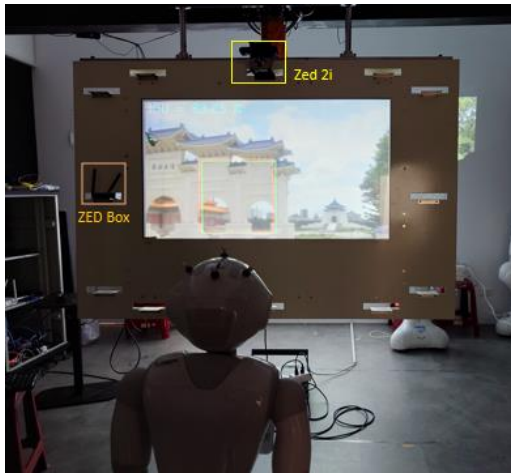


Figure 5. The design flow chart of the information fusion.

As a result, the whole design process is presented in Figure 5. First of all, the necessary parameters would be measured in advance and set as initial. Next, all the results of each device would be transferred to a unified coordinate to build up the integrated system, and then, the result of the camera will be recalibrated to fit the realistic situation. After calibrating the devices, the spatial locations of the information fusion would be computed, and also, the size of the bounding box would be determined by the distance between the observer, the display, and the object. The spatial location of the display, especially, will be next computed, which will determine the relationship of the display and the information. Finally, transfer the spatial location to the pixel on the display finishes the information task.

## 3. Experiments and discussion

In this section, the implementation of the information fusion and the calibration result of rotating the camera is discussed. The configuration is shown in Figure 6. The task is running on ZED Box which is powered by Jetson Orin NX 16 GB. The reference observer is set 2 meters away from the display with a humanoid robot, and the object is set on the opposite side of the display, which is projected on a white wall 1 meter away. The frame of the display reserves several erection points for the camera to build a flexible system. In this experiment, both camera is set at the top-center of the front and the back to present an intuitive and general case of an information fusion.



**Figure 6.** The experimental configuration of the information fusion task.

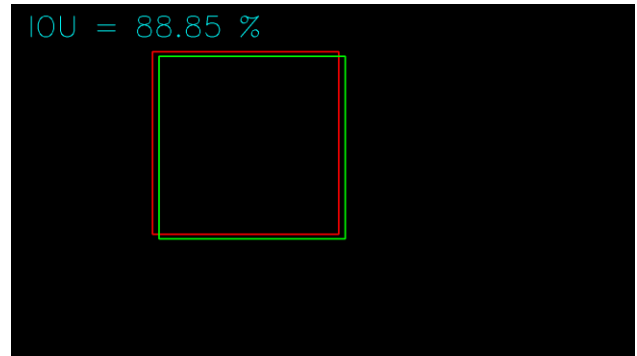
To judge the accuracy of the information fusion, the spirit of the intersection over union is imported to show how accurate the result is. The location and the size of the bounding box is adopted to show the relationship of the desired and actual content location visually, and the ratio of the overlap part and the union part is calculated as the result.

Figure 7 shows some case results of the calibration experiment which is outputted on the transparent screen. The black background corresponds to the transparent in a transparent display in practice. The red box represents the ground truth where the reference location of the observer is selected manually and the object location is detected by the camera. The green box shows the calibrated result, and the location of the box follows the detected result of the observer and the object. The benefit of designing the validation method in the form of computing the overlap region is that the accuracy only depends on the location of the observer. When the object is detected, the shape of the information box is also decided, and the size of each box is calculated by the object detect result and the ratio of the vertical distance of the display to the object and the observer to the object.

If the two boxes overlap completely, which means the observer is detected in the ideal location. This can easily reflect the error caused by the camera on the display for validation, and analyze how to fix the error.

In this experiment, comparing the results of camera calibration and the results before calibration will not be discussed, and only the camera-calibrated information fusion result will be shown. Although the calibration could offer an acceptable result, there are still some errors that exist, and the result also depends on the distance of the observer and the object to the display.

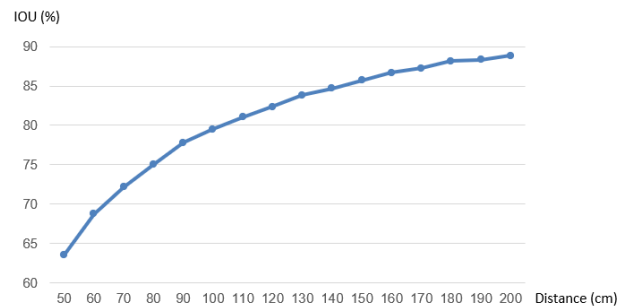
Table 1 shows the result of the accuracy changes when observing from the region of 50 cm to 200 cm away from the display and records the results every 10 cm.



**Figure 7.** The content of the calibration experiment. (Distance to object = 100, Distance to observer = 200)

**Table 1. Experiment results for different observer distance (distance to object = 100)**

Distance	IOU	Distance	IOU	Distance	IOU
50	63.52	100	79.52	150	85.73
60	68.75	110	81.02	160	86.69
70	72.18	120	82.37	170	87.28
80	75.01	130	83.85	180	88.14
90	77.77	140	84.69	190	88.35



**Figure 8.** Trend Chart for different observer distance (distance to object = 100)

From the table and Figure 8, we can see that as the observer distance is shortened, the accuracy falls, and if the distance of the display to the observer is less than the distance to the object, the accuracy falls greatly. Therefore, if the application of the information fusion task requires a short distance of the observer to the screen. It would be worth applying a precise calibration method to have better results.

#### 4. Conclusion and future work

A Modular information fusion building method and the key parameters have been presented, which enables the construction more conveniently, without the need for extensive camera position calibration. Through simple experiments, the effect of the calibration has been reflected in the accuracy of the fusion location for different distances. In the future, we will continue to make efforts on modularizing the information fusion system and let the module be extended to any placement degree or different type of display, such as curved displays or irregular shape displays.

**5. Reference**

[1] Hirakawa M, Koike S. A Collaborative Augmented Reality System Using Transparent Display. IEEE Sixth International Symposium on Multimedia Software Engineering [Internet]. 2025 [cited 2025 Mar 31];410–6. Available from: <https://ieeexplore.ieee.org/document/1376689>

[2] Costanza E, Kunz A, Fjeld M. Mixed Reality: A Survey. Lecture Notes in Computer Science. 2009 Mar 27;47–68.