A Multi-Detailed Spatial Immersive Display

Tsuyoshi Minakawa† **Toshio Moriya**† ‡ **Masami Yamasaki**† **Haruo Takeda**† ‡

†Systems Development Laboratory, Hitachi, Ltd.
1099 Ohzenji, Asao-ku, Kawasaki, Kanagawa 215-0013, JAPAN
‡Nara Research Center, Telecommunications Advancement Organization of Japan
8916-19 Takayama, Ikoma, Nara 630-0101, JAPAN

{*mint, moriya, yamasaki, takeda*} *@sdl.hitachi.co.jp*

Abstract

A spatial immersive display that demonstrated non-uniform pixel density is introduced. The non-uniformity is determined by using the following concept — "*detailed images for important parts, rough images for the others.*" To realize a system based on the concept, we used multiple projectors arranged non-uniformly.

First, we created a immersive display system using four screens with four projectors. Then, we prepared a motion ride to compel the user to gaze at a part of the front screen, and to estimate their viewpoint and posture without their wearing any tracking devices. Finally, we added two more projectors to provide detailed images at the center of the front screen. After geometric and color calibration using a digital still camera, seamless images of multiple detail were observed on the front screen.

Key Words: immersive display, multi-detailed image, multi-projection technology

1 Introduction

The introduction of CAVE$^{\text{TM}}$[2] in 1993 spawned the proposals of a lot of other spatial immersive display systems. These newly proposed systems adopted different approaches to bring about an improvement in the emerging technology. For example, some displays increased the number of screens to activate a larger viewing angle [6][14], and others modified the shapes of screens efficiently to cover a large field of view [5][7][8].

On the other hand, a lot of multi-detailed techniques, such as display systems [3][15], rendering methods [9], and evaluations [10][13], were also proposed. Furthermore, two multi-detailed stereoscopic video systems — one was compositing images in optical way, and the other was in digital — were introduced in [4] with an input device of multi-detailed images. The objective of all of these techniques was to provide a detailed image near the gazing point, thus almost of these type of techniques take an eye-tracking approach.

We propose in this paper a multi-detailed spatial immersive display that embodies many of the advantages of both types of displays. We introduce our approaches to improvements of a simple surround-screen display system in Section 2, describe our hardware environment and its calibration method in Section 3 and Section 4, respectively. And finally, we show some experimental results in Section 5.

2 Our Approach

In this section, we describe our approaches to making a multi-detailed spatial immersive display. As an example, a snapshot of our final environment and its system block diagram are shown in Figure 1 and Figure 2, respectively.

First, we chose a CAVE$^{\text{TM}}$-like immersive display as the basis of our system concept. Among spatial immersive displays, this type of display is easy to construct and render images because the shape of the screens is rectangular. This feature allows us to make an inexpensive system and to achieve a higher frame-rate in realtime rendering. In the comparison with head-mounted displays (HMDs), it also has some advantages, such as higher resolution, larger viewing angle, and saving lag when the user turns their head.

Next, we integrated a motion-ride system that has four degrees of freedom (DOF) — roll and pitch as rotation, up and down, forward and backward as translation — to the space enclosed by the screens [11]. This modification has a great disadvantage in that it prohibits the user from wandering around the display freely, but it brings us the following advantages in return.

- Of course, the ride delivers the sensation of acceleration to the user, thereby improving the immersive feelings of being ensconced in a virtual world.
- In a multi-screen system, it is very important to adopt a head-tracking technique to correctly display images to enhance the user's experience. Thus, we deal with the user's view in a real-world and in a virtual-world independently, and use the motion-ride system as the head tracker. We calculate the head position of the user from the position and posture of the ride (or from a series of motion data [11], or in consideration of some other psychological factors) by using a model of the human body. So as to not spoil the fun of the user, it is desirable

to decrease the lag of the tracking [1]. This way of head tracking allows us a lag of few frames, though it sacrifices the accuracy of the estimated result. It also allows us to pre-render high-quality images in non- or semi-interactive content in which the view frustums used for rendering have been previously determined for some patterns. And, in this way of tracking, the user needs to wear glasses only when they want to view a stereoscopic image.

- The ride compels the user to gaze at a part of the front screen, though this must be verified by an future experiment. It allows us to consider the left, right, and bottom screens as peripheral-view areas, and the top and rear screens as meaningless ones. The small number of screens allows us to use the remaining projectors for other purposes. For example, when six projectors are available similarly to the display introduced in [14], we can use one for every screen, and the remaining two are dedicated to the front screen to provide the detailed images. Moreover, it is unnecessary to prepare data for the whole viewing angle, thereby, deriving benefits, such as, easing the task of designing of input devices, and reducing the total amount of data.

Finally, in order to provide more detailed information to the user, we add more projectors to the system, especially for the front screen, which is where the user will customarily being gazing. The projectors are roughly arranged, and then calibrated with a digital still camera so that their projections can form a seamless projection with the projection of the base one. This approach to showing multi-detailed images has the following advantages and disadvantages.

Advantages

- We can control pixel density on the screen easily by varying the distance between the front screen and the extra projectors in the in-focus range. We can increase the number of projectors assigned to the detailed images if needed.
- It is easy to create stereoscopic images with polarizing filters by stacking two projections because we can calibrate the correspondence between those two projections accurately in consideration of the distortions presented by the screen and the projector lenses.
- In this paper, we introduce a two-leveled type of detail projection, however, it is easy to increase the level of detail projection, in similar way. The only constraint is that each projector must not cast shadows with the lights of the other projectors.
- Indeed the system does require calibration, however, we consider that task simple enough in comparison with the setup required with the multi-detailed HMDs which composite images in optical way.
- When the system is calibrated for static-type compositing (described in Section 4.2 and shown in Figure 3(a)), the detailed projections and the rough projection form a seamless projection by edge blending. The user will not perceive the boundary between the two levels of detail.
- When the system is calibrated for dynamic-type compositing (described in Section 4.3 and shown in Figure 3(b)), the detailed insets can take arbitrary shapes in the detailed projection area, though the system needs the assistance of content-player software. This is because the two levels of detail projections coexist in the detailed projection area.

Disadvantages

- The system provides detailed images to an area for which only rough images are needed, because it does not take into consideration the current gazing point and gaze direction of the user.
- When the system is calibrated for static-type compositing, we cannot change the positions of the detailed projections. It is because the system must be re-calibrated previously, and that cannot be done in realtime.
- When the system is calibrated for dynamic-type compositing, the boundary contour between two levels of detail are either glittery or blackened because the two levels of detail have pixels of different sizes that are usually not stacked in perfect arrangements, though we can try to use anti-aliasing software to reduce this type of error.

We consider that the most suitable application for our system is a one-player game, such as a driving simulator or a roller-coaster game, especially non- or semi-interactive content which can use pre-rendered high-quality images.

3 System Components

We introduce here the total environment of our system consisting of four screens and six projectors.

First, we made a spatial immersive display using four screens and four projection units. The bottom screen measures 2.4 by 0.8 meters, and the others are 2.4 by 1.8 meters. Each projection unit consists of three components — a digital projector, an image-modifying unit, and an image-providing PC.

Each projector can be roughly set up under the condition that its projection area covers the whole area of each screen. Of course, similar to HMDs [12], we need to perform geometric calibration to make the image-projected area exactly the same for the whole screen area in consideration of the arrangement errors and the distortions created by the screen and the projector lenses. However, we leave this task to the image-modifying unit. The images handled by the projectors have a resolution of 640 by 480 pixels.

Accurate modification of the system geometry is done in realtime using the image-modifying unit. The unit, which was designed and developed by us, can make geometric and color modifications to the provided images. The parameters used in these pixel-wise modifications were already made and stored in the memory of the unit. The unit can handle provided images of up to 1024 by 768 pixels at the refresh rate of 60 Hz.

The images provided by the PCs are previously rendered movies and/or realtime-rendered interactive content. They are run on our content viewer so that images rendered for each screen can be synchronized with each other.

Next, we added a motion-ride effect (with a joystick as a user interface) to the space enclosed by the screens. It has four DOF, and is controlled by the motion-controller PC to be in sync with the displayed images. As a result, the user feels a natural sensation of accelerating, and it also enables us to estimate the user's view.

Finally, we prepared two more projection units to complete an immersive environment for the projection of detailed images. We assign the inner area of the front screen, where users will mainly focus their attention, to the projectors of the additional units. The projections of these projectors are usually brighter than the base one, because they cover smaller area than the base one does. In order to absorb this difference of the brightness, we use neutral-density(ND) filters of suitable attenuation level, even though we can do so with the image-modifying unit itself. We do this because we want to maintain as much as possible the contrast and the color resolution of the original images in order to provide high-quality images. The images handled by the additional projectors have a resolution of 1024 by 768 pixels.

The precise calibration of the system described above (i.e., to set the parameters to be stored in the image-modifying unit's memory), we use a digital still camera controlled by a PC. It is desirable that the camera have a variable-speed shutter and that it can be controlled adaptively, because we need accurate data in color calibration. The translation functions from pixel value to luminosity are calibrated beforehand under a good condition.

4 System Calibration

In order to provide seamless images with the above system, we need to make geometric and color calibrations. There are two types of compositing, static-type and dynamic-type, for providing multi-detailed images. In this section, we explain the difference between those two types. Then, we describe in detail how to calibrate the system for the static-type compositing. And lastly, we touch on the calibration method for the dynamic-type briefly.

4.1 Two ways of compositing

The process of static-type compositing and that of dynamic-type are shown in Figures 3(a) and 3(b), respectively.

The static-type compositing always uses detailed images for the stacked-projection area. This type of compositing can be handled by the image-modifying units themselves, so that we can realize the displaying of multi-detailed content at a high frame-rate. A special feature of this type of compositing is that the transition of the detail is made smoothly enough for no one to notice it happening. It gives the illusion that the additional projectors are making small contributions, but the great amount of information being unobtrusively provided in the detailed part by those projectors will make users naturally become more immersed.

In contrast to that, the dynamic-type compositing uses detailed images for only the important areas. It needs the assistance of compositing software in addition to the image-modifying units, because the parameters stored in the hardware are static while the important parts of an image will be varying dynamically. This type of compositing has some advantages over the static-type compositing. For example, it is easy to call the user's attention to the important parts. And other example is that it needs less image data than the static-type compositing — this is a great advantage especially in the use of telecommunicating applications. Of course, there are not only advantages, there are some disadvantages that have to be coped with. For example, it is necessary to take care that the synchronization among the projected images is precise, because the transition of the detail occurs suddenly. In other words, a loose synchronization will make the boundary area between the rough portion and the detailed portion look black-colored or glittered. Moreover, as we have already mentioned in Section 2, when you use digital projectors, you must compensate mentally for the anti-aliasing on the boundary area.

4.2 Calibration for the static-type

In this subsection, we describe how to calibrate the system for the static-type compositing.

We use vertical and horizontal gray codes to make correspondence maps between the coordinates of the camera and that of the projectors. In addition to that, we use other stripe codes to make the maps more precise. Then, we make geometric-modification parameters for the stacked projections from the maps. (That parameter for the base projection should be fixed, because it has already been calibrated in relation to the front screen in a similar way, in consideration of the adjacent projections.)

For color calibration, we prepare characteristic functions for each projection unit, and decide their parameters using some additional patterns. We perform edge-blending between the stacked projections first, and then, do the same between the base and the stacked projections. To prevent false ridges near the corners of the stacked projections, we perform an elliptic blending for the second edge-blending, which has an elliptic contour of luminosity at each corner of the stacked projections.

Of course, if there were no error, any blending method brings about perfectly seamless images on the screen as long as it is a theoretically correct one. However, there are some errors that when present thwart attempts at perfectly theoretical optical compositing. For example, there is measuremental error, of course. Another example is discrete errors of luminosity — for each color, luminosity is processed in 8 bits for the input and for the output, though it is processed in 10 or more bits for the internal processes. One more example is that there are errors caused by the size of pixels. Both levels of detail pixels have certain sizes and they usually are not stacked in perfect arrangement, so that errors arise for anti-aliasing.

4.3 Calibration for the dynamic-type

We briefly explain now how to calibrate the system for the dynamic-type compositing.

In this type of compositing, the requirements for the modification hardware are to make the geometry of the stacked projection units consistent with that of the base projection unit, and make the color characteristic functions of the stacked projection units exactly the same as that of the base one.[1] These requirements are easily accomplished because they are the same as those of the calibration method described in Section 4.2 except that it is not necessary to blend the boundary between the stacked projection area and the base one. As a result of that, the base projection unit needs no additional modifications to convert the system into multi-detailed one.[1]

5 Experimental Results

The experimental results for the blending are shown in Figure 4. Out of the two types of compositing, we calibrated the system for the static-type.

When we observed the calibrated images, we made the two following observations: First, the system provided a seamless image after geometric and color calibrations. Secondly, the system provided rich information to the area where users mainly focus their attention. We could recognize letters in the detailed-image area while we could not do so in the rough-image area.

Next, we show the results of the comparison between elliptic and rectangular blending in Figure 5. From this figure, we can see that the elliptic blending was working well while the rectangular blending was not performing as well. Owing to the elliptic blending, no false ridge is observed. This indicates that our way of blending is apparently one good solution to get seamless images of high quality.

6 Concluding Remarks

6.1 Conclusion

We have developed a multi-detailed spatial immersive display. It provides detailed images of important parts and rough images of the others by dint of the non-uniform arrangement of projectors. In order to keep a user's attention directed at the detailed projection area, a motion ride is used to control the user's viewpoint and posture. Its geometric and color calibrations are made using a digital still camera and some calibration patterns. After the calibrations, we obtained multi-detailed seamless images in realtime using image-modifying units whose memories contained the calibrated parameters. We found that, in at least the closely-observed area of a user, the system can provide more information to the user than any other immersive displays using the same resources. Owing to these results, we consider that our system is a significant improvement over the current spatial immersive display system.

6.2 Further Works

As to our further works, we will explore the following two points. One is to improve the quality of compositing. In order to that, we are trying to use a more precise optical model of the projectors, though it will take longer for calibration. The other is to realize a real-time dynamic-type compositing, especially an optical compositing of video and CG images. We will try to composite detailed video images made using the chroma-key technique into realtime-rendered background CG images. We consider that this approach is an effective one for use in a telecommunication setup of limited resources.

References

[1] R. S. Allison, L. R. Harris, M. Jenkin, U. Jasiobedzka, J. E. Zacher: Tolerance of Temporal Delay in Virtual Environments, *Proc. of IEEE Virtual Reality 2001*, Yokohama, 2001, pp. 247–254.

[2] C. Cruz-Neira, D. J. Sandin, T. A. DeFanti: Surround-Screen Projection-Based Virtual Reality: the Design and Implementation of the CAVE, *Proc. of ACM SIGGRAPH'93*, 1993, pp. 135–142.

[3] W. S. Geisler, J. S. Perry: Variable-Resolution Displays for Visual Communication and Simulation, *The Society for Information Display*, Vol. 30, 1999, pp. 420–423.

[4] K. Goshi, K. Matsunaga, H. Nagata, K. Shidoji, H. Matsugashita: Digital Stereoscopic Video System with Embedded High Resolution Images, *Proc. of IEEE Virtual Reality 2001*, Yokohama, 2001, pp. 191–197.

[5] W. Hashimoto, H. Iwata: Surrounded Spherical Immersive Display with Stereoscopic View, *Proc. of the 6th VRSJ Annual Conference*, 2001, pp. 355–358. (*in Japanese*)

[6] M. Hirose, T. Ogi, S. Ishiwata, T. Yamada: Development and Evaluation of Immersive Multi-screen Display "CABIN," *IEICE Transactions*, J81-D-II, No. 5, 1998, pp. 888–896. (*in Japanese*)

---

[1] These statements are valid only when the *black* level of each projection is dark enough.

[7] H. Iwata: Closed Surface Display by Rhombic Dodecahedron Screen, *Proc. of the 1st VRSJ Annual Conference*, 1996, pp. 91–94. (*in Japanese*)

[8] H. Iwata: Seamless Surrounded Spherical Display, *Proc. of the 3rd VRSJ Annual Conference*, 1998, pp. 155–158. (*in Japanese*)

[9] T. Ohshima, H. Yamamoto, H. Tamura: Gaze-Directed Adaptive Rendering for Interacting with Virtual Space, *Proc. of IEEE Virtual Reality Annual International Symposium '96*, Santa Clara, 1996, pp. 103–110.

[10] D. Parkhurst, E. Culurciello, E. Niebur: Evaluating Variable Resolution Displays with Visual Search: Task Performance and Eye Movements, *Proc. of the Eye Tracking Research and Applications Symposium*, Palm Beach Gardens, 2000, pp. 105–109.

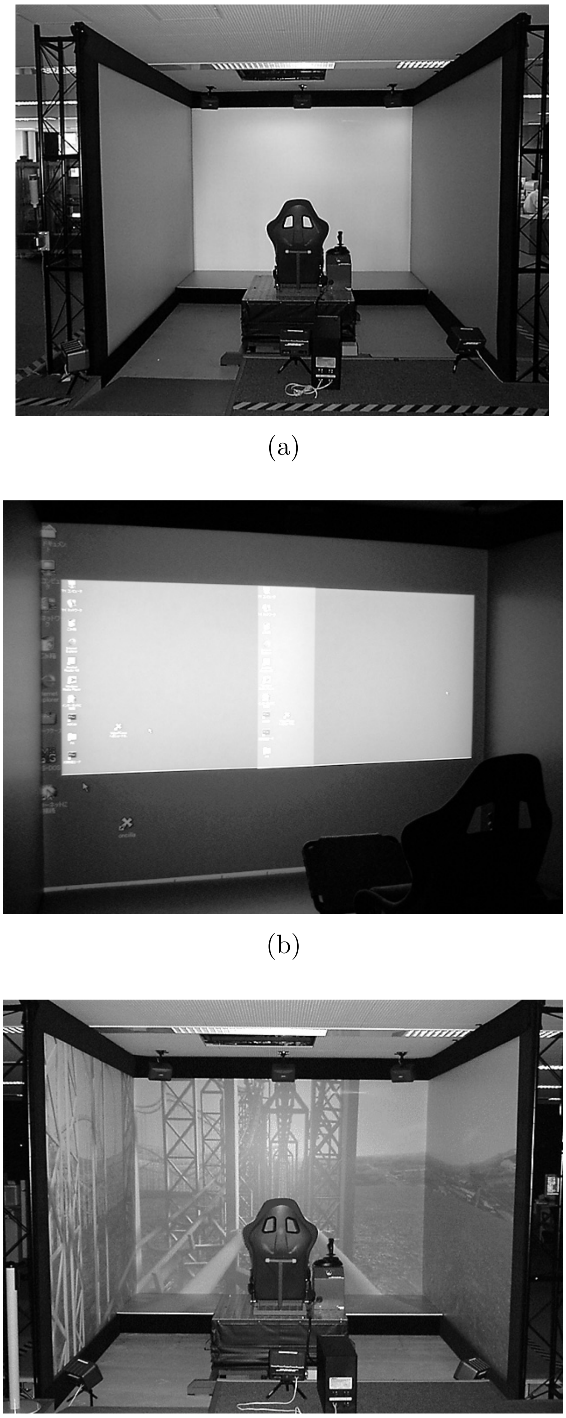
[11] H. Takeda, S. Kiyohara, K. Chihara, H. Kawase, Y. Matsuda, M. Yamasaki: Multi Screen Environment with a Motion Base, *Lecture Notes in Artificial Intelligence*, Vol. 1834, Springer-Verlag, Berlin Heidelberg, 2000, pp. 303–312.

[12] B. A. Watson, L. F. Hodges: Using Texture Maps to Correct for Optical Distortion in Head-Mounted Displays, *Proc. of IEEE Virtual Reality Annual International Symposium '95*, Research Triangle Park, 1995, pp. 172–178.

[13] B. Watson, N. Walker, L. F. Hodges: Managing Level of Detail through Head-Tracked Peripheral Degradation: A Model and Resulting Design Principles, *Proc. of ACM Virtual Reality Software and Technology '97*, Lausanne, 1997, pp. 59–64.

[14] T. Yamada, M. Hirose, Y. Iida: Development of Complete Immersive Display: COSMOS, *Proc. of VSMM98*, Vol. 2, 1998, pp. 522–527.

[15] A. Yoshida, J. P. Rolland, J. H. Reif: Design and Applications of a High-Resolution Insert Head-Mounted-Display, *Proc. of IEEE Virtual Reality Annual International Symposium '95*, Research Triangle Park, 1995, pp. 84–93.



(a)

(b)

(c)

Figure 1. Our experimental environment. (a) Snapshot. (b) Initial projections on the front screen. Without ND filters, the stacked projections are about 15 times brighter than the base one. (c) Typical application.

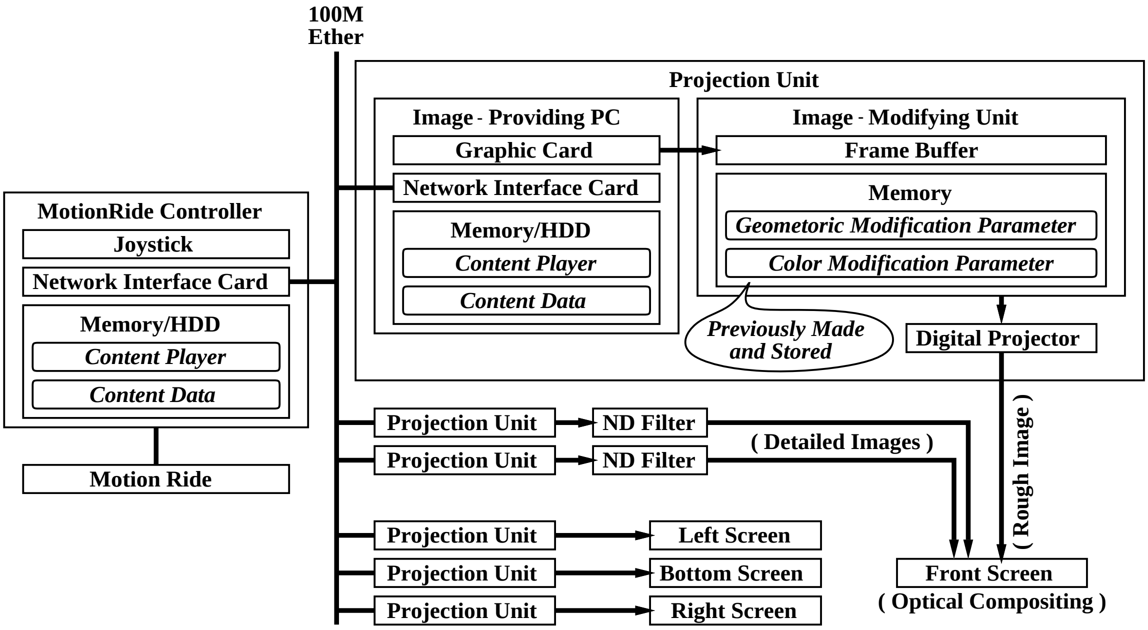
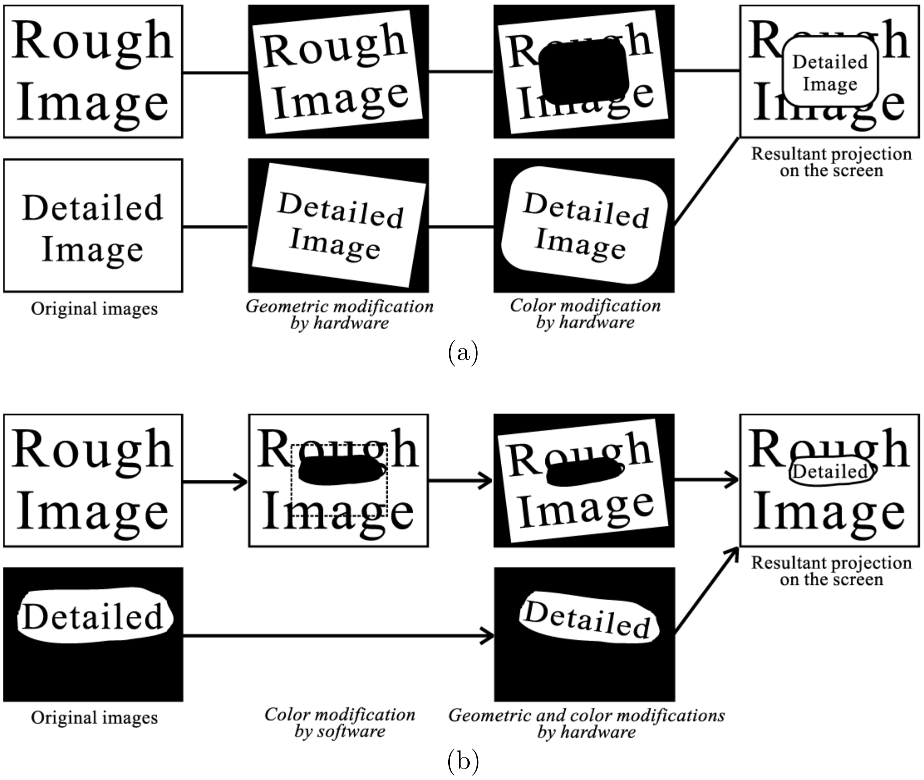


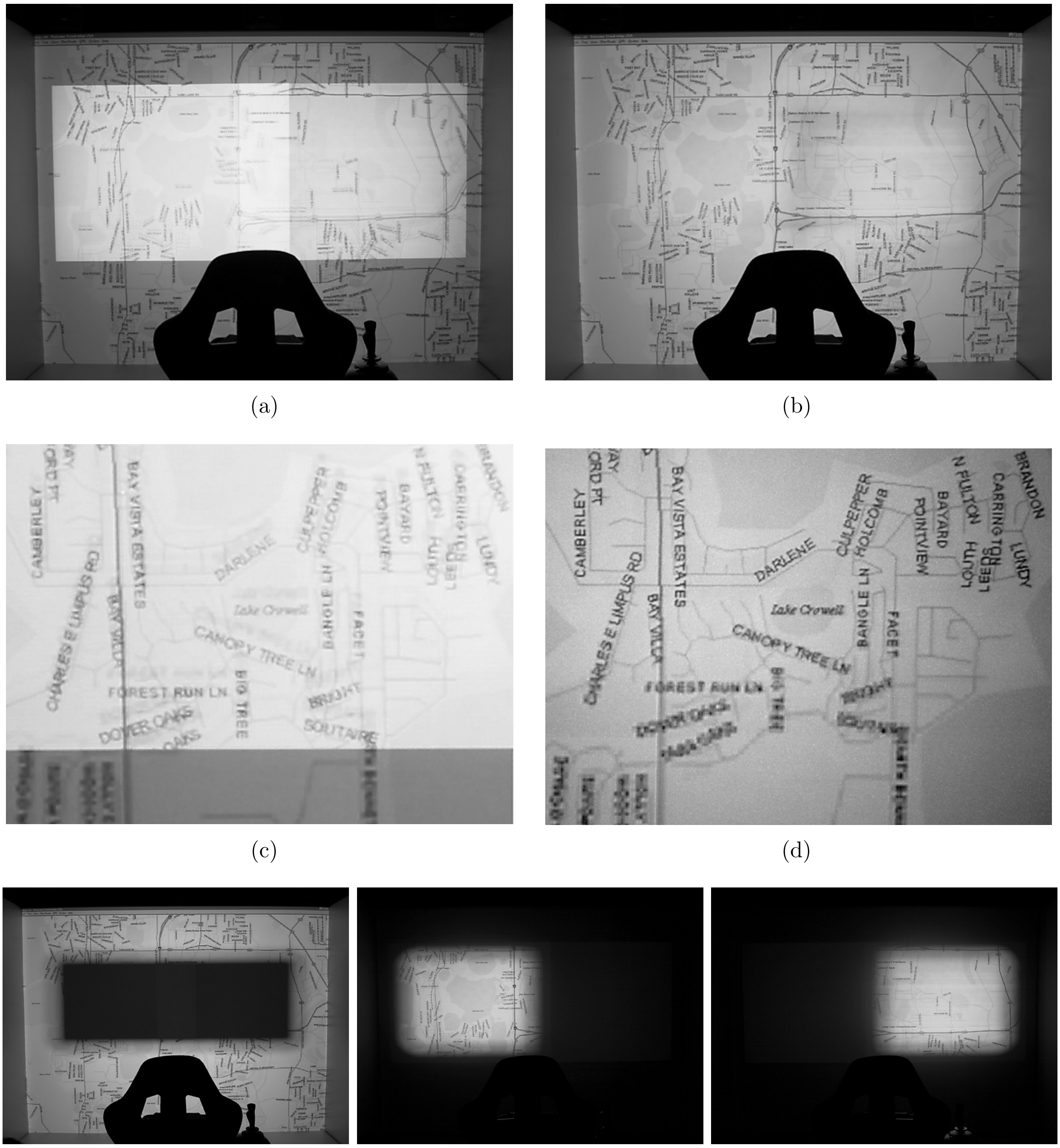
Figure 2. System block diagram.



(a)

(b)

Figure 3. Data flow diagram to provide multi-detailed images. (a) The static-type compositing. This type of compositing can be handled by the image-modifying units themselves, though the position of the detailed area is unchangeable. At the color-modification phase, an elliptic blending is performed. (b) The dynamic-type compositing. The system needs the assistance of content-player software, though the detailed insets can take arbitrary shapes in the detailed projection area.

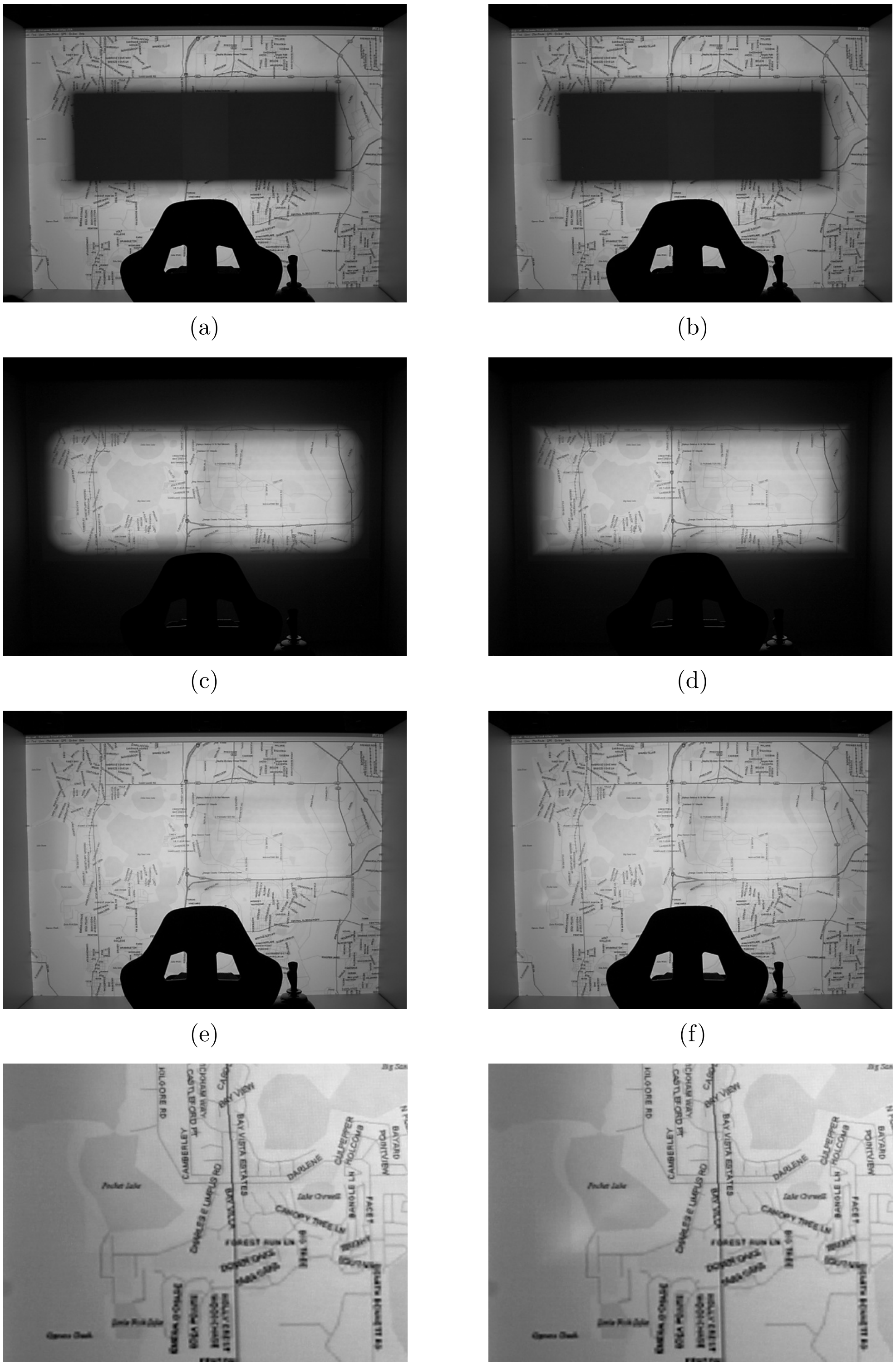


(a) (b) (c) (d) (e) (f) (g)

Figure 4. Elliptic blending.[2] (a) Initial state of the projections on the front screen. (b) After calibration. (c) Enlargement of a boundary area at the initial state. The detailed area is brighter than the rough area, and some words are projected double. (d) Enlargement of a boundary area after calibration. The words in the detailed-image area are legible for the most part while the words in the rough-image area are quite illegible. (e)(f)(g) Projection of each projector on the elliptic blending.

---

[2] We used the 1999/2000 INFOCOMM International® Projection Shoot-Out® Software to evaluate our system's capabilities.



(a) (b)

(c) (d)

(e) (f)

(g) (h)

Figure 5. Comparison between elliptic and rectangular blending. Lefthand and righthand images are the results of elliptic and rectangular blending, respectively. (a)(b) Projection of the rough image. (c)(d) Projection of the detailed image. The ridges, which will be false ridges when optical compositing, are observed near the corners in the rectangular blending while they are not in the elliptic blending. (e)(f) Composited image. (g)(h) Enlargement of the region near the bottom-left corner of the detailed area. The corner of the detailed area is perceived in the rectangular blending while it is not in the elliptic blending.