CoGAME: Manipulation Using a Handheld Projector

Kazuhiro Hosoi* Vinh Ninh Dao† Akihiro Mori‡ Masanori Sugimoto§
Interaction Technology Laboratory
Department of Frontier Informatics, The University of Tokyo

Abstract

An example of an application enhanced by the ”manipulation-by-projection” technique, this cooperative game allows players to visually and intuitively control a robot with projectors. Players interchangeably move and connect their projected images to create a path that leads the robot to its goal.

1 Introduction

Due to the development of recent electronics technologies, projectors become not only smaller but also cheaper and power-hungry. In fact, Siemens' researchers have developed a mobile phone featuring a built-in projector system [Siemens]. Moreover, palm-sized and battery-powered commercial projectors[Toshiba 2006], which use low-power LED lighting are already available.

If a projector is mounted onto a mobile device, it can create a larger information display without external equipment, by projecting images onto surfaces such as walls, the ceiling, or the floor. Moreover, as we can freely move such a small and lightweight projector in a 3D space, we can use it not only as an output device but also as an input device, for example, manipulating real movable objects such as robots. Therefore, we have explored possibilities of human-robot interaction techniques via image projection.

In this paper, we propose the "manipulation by projection" technique that allows users to visually and intuitively control a robot using their projectors. By projecting a variety of information including the robot's state, its surroundings, and human commands to the robot via a handheld projector, users can confirm their commands and the reactions of the robot. Although a user could perform a similar task by using a Head-Mounted Display (HMD), the proposed technique also allows other users to see the information, and these users can interact with the robot without conflict. Furthermore, our proposed technique enables a simple robot with only a few sensors and interfaces, to act as though it were intelligent. Instead of using sensors attached to the robot, a mobile device can recognize the robot's state and its surrounding environment through a camera attached to the device. The system generates a robot motion plan, and the robot behaves as intelligently as if it had adapted to the environment by itself.

The proposed technique will be usable in a wide range of applications in a future society where a variety of robots, including simple robots, will be part of our daily lives. Suppose that we want to rearrange several pieces of furniture in an office or a laboratory. We may plan their arrangement, relocate the furniture, and evaluate the arrangement via discussion amongst ourselves. We can use the proposed technique for this task, by visually and accurately controlling robots that relocate the furniture. The user projects an image of the furniture layout into the real environment, and the robot accurately carries the furniture by following the projected image. While the robot is working, other users can check the new layout and discuss it.

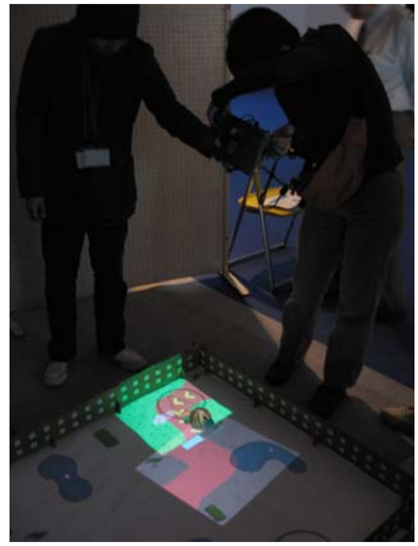

Figure 1: *CoGAME: A robot navigation game using projected images given through mobile devices*

In an entertainment scenario, the proposed technique can enhance playing with toys, such as small robots and radio-controlled cars, by using Augmented Reality (AR) technologies. The player can make the toys act in response to virtual objects in the projected display, and can intuitively manipulate the toys by moving the projected display. For a multi-player game, players can cooperatively play with the toys by connecting or sequencing the images. Moreover, by displaying and sharing the manipulation, the proposed game enhanced by our "manipulation by projection" technique is able to amuse not only the players, but also the game's spectators. Furthermore, as no special devices are required to be installed in the environment, players are able to enjoy the game anywhere, without being concerned about any limitations of the game field in the real world. Thus, our proposed technique can be used as a novel mobile entertainment tool that can integrate a virtual world with the real world anywhere.

As an example of the above application, we have developed an interactive game called CoGAME. The idea of this game is to lead a turtle robot to a destination, using a handheld projector. By projecting part of a road image onto the floor, the robot can walk along the road as shown in Figure 1. The players take turns to move and connect their own projected road image, thereby creating the path that leads the turtle to the goal.

2 Related Works

Some researchers have proposed interactive systems using handheld projectors [Raskar et al. 2003; Miyahara et al. 2005; Cao and Balakrishnan 2006]. They focus on interaction techniques, for ex-

*e-mail: hosoi@itl.t.u-tokyo.ac.jp
†e-mail: dao@itl.t.u-tokyo.ac.jp
‡e-mail: mori@itl.t.u-tokyo.ac.jp
§e-mail: sugi@itl.t.u-tokyo.ac.jp

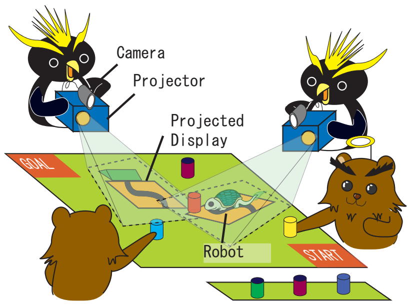

Figure 2: *CoGAME overview*

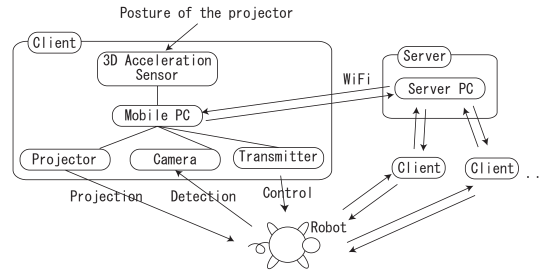

Figure 3: *System configuration*

ample, manipulations with fingers on a projector display [Miyahara et al. 2005], stabilizing projected images [Raskar et al. 2003], and so on.

Other researchers have mainly focused on the manipulation of virtual information. Furthermore, they have yet to develop multi-user applications using handheld projectors. Our system is for a multi-user applications that enables users to manipulate not only virtual information but also real robots.

AR technologies are used to display information intuitively by overlaying virtual objects in the real world. There are many applications of AR technologies, including a system for supporting collaborative work. AR technologies have also been applied to robotic interface and human-robot interaction studies. *Virtual Humanoid* [Shoji et al. 2006] is an AR system that synchronizes a humanoid robot and a virtual avatar using an HMD. Users can experience handshakes with a computer-graphic avatar. This system integrates sensing and control systems to keep the robot and the avatar motions synchronized during their physical contacts with humans. Augmented Coliseum [Sugimoto et al. 2005] is a game environment using an AR technology that enables a player to play a game using small robots with virtual functions. Projecting images onto the areas that correspond to the actual positions and directions of the robots can enhance playing with them.

The above systems enables visualization of the robot information using a display, an HMD, or a projector embedded in an environment. Because there is neither need of extra equipment for the robot, nor specific environment, our proposed system allows a user to control robots anywhere.

3 How to play

As shown in Figure 2, players hold projector-enabled devices in their hands, and project images onto the floor. When one player, *player A* projects an image at the robot, it moves along the road in the projected image. (The required direction for the robot is displayed in the image.) When the robot reaches the end of the displayed road, it stops. At this point, another player, *player B*, projects an image near *player A*'s image, and tries to connect the road to *player A*'s road by adjusting the projected image. When the robot has reached the end of *player A*'s road, and the road is connected to *player B*'s road, the robot can then move onto *player B*'s road. After the robot leaves *player A*'s road, the CoGAME system randomly changes *player A*'s image. *Player A* then connects the road in this new projected image to the end of *player B*'s road. If there are more than two players, they take turns to connect their displayed road images, thereby leading the robot to the destination.

It is possible that *player A* alone couled lead a robot to the destination without the help of *player B*, by moving his projected image in the direction of robot movement before the robot reaches the end of the road in the original displayed image. In order to prevent this operation for retaining the feature of CoGAME, that is, a cooperative game, we made the following rule: each player has control of the robot only for a limited time, and must lead the robot into another player's road within that time.

4 System

The CoGAME system comprises multiple clients, a server, and a turtle robot. As shown in Figure 3, each client has a projector (Toshiba, TDP-FF1A, 800x600 resolution), a mobile PC (Sony Vaio Type U, CPU: Intel Core Solo U1400 (1.2 GHz), memory: 512 MB), a USB camera (Logicool Inc. Qcam for Notebooks Pro, 640x480 pixels), a three dimensional (3D) acceleration sensor (Sunhayato Inc, MM-2860) and a transmitter. The transmitter is used for sending commands to the robot via infrared communication. The client package's total weight is about 1 kg, and players can easily move it around. The client produces a projected image, detects the robot, and controls it. We use a laptop PC (CPU: Pentium M 1.8 GHz, memory: 1 GB) as the server. The server manages all clients and generates the robot motion during game play. Data communication between the server and the clients is via a wireless LAN. The robot only has a receiver and an actuator, thereby remaining simple and small.

The software architecture of CoGAME is shown in Figure 4. First, each client recognizes the pose of its projector by retrieving the data from its 3D acceleration sensor. Using the pose of the projector, the client calculates the transformation between the source image and the projected image. At the same time, it detects the robot from the camera images and estimates the position and the direction of the robot. The client translates the data from the camera image plane to a corrected image plane by using the information retrieved from the tilt sensor. The server gathers and analyzes the robot data, and decides which client the robot should follow. Then the server generates a robot motion plan and sends it to the client who has the authority for robot control at that time.

4.1 Projection correction

The shape of the projected display depends on the projection angle and the distance between the projector and the screen surface. Consequently, projected images are distorted unless projected orthogonally to the screen. This distortion raises the level of the human user's cognitive load in understanding the image content. Therefore, it is necessary for the system to correct the shape of a projected

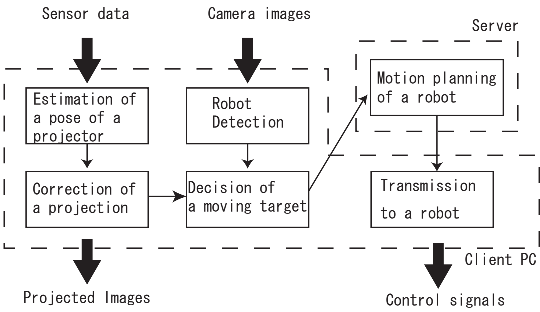

Figure 4: *Software Architecture*

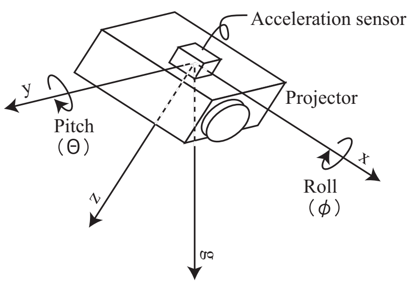

Figure 5: *A pose of a projector*

image by estimating the distortion in the projection.

Our approach to correcting the projected images is to use a 3D acceleration sensor as a tilt sensor to retrieve the roll angle and the pitch angle of the projector. When a player projects an image to a plane which is parallel to the ground plane, the projected image depends on the roll angle and the pitch angle of the projector as shown in Figure 5.

Consider a point $A(x_A, y_A)$ in the source image. This point is projected to a point $a(x_a, y_a)$ on the projected display by a perspective transformation. We can calculate this transformation through a series of rotation transformations around the vertical and horizontal axes. Figure 6 shows that the following equation changes a rectangle *ABCD* in the source image to a quadrangle *abcd* in the projected image.

$$\begin{pmatrix} x_a \\ y_a \end{pmatrix} = \mathbf{P}(\theta)\left(\mathbf{R}(\phi)\begin{pmatrix} x_A \\ y_A \end{pmatrix}\right) \quad (1)$$

$$\begin{pmatrix} x_a \\ y_a \end{pmatrix} = \mathbf{V}(\theta,\phi)\begin{pmatrix} x_A \\ y_A \end{pmatrix} \quad (2)$$

Here, ϕ and θ are angles around the X-axis and Y-axis of the tilt sensor, and P and R are the projection matrix and the rotation matrix, respectively.

The system selects a projected image area, which is a rectangle *mnpq* within the projected display *abcd*. By using the inversion matrix of the V matrix in Eq.(2), the system transforms the rectangle *mnpq* in the projected image to the quadrangle *MNPQ* on a source image. Then, the system calculates a transformation, called a homographic transformation, between the rectangle *ABCD* and the quadrangle *MNPQ* by using the four pairs of corresponding vertices in these quadrangles. In this way, the system transforms a source image to a corrected image using the homographic transformation.

To display an entire source image, the corrected image has to be smaller than the projected display's size. Therefore, we need to de-

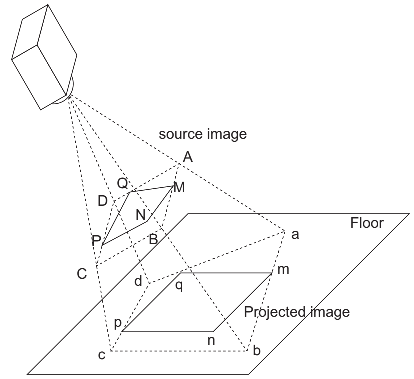

Figure 6: *Homography*

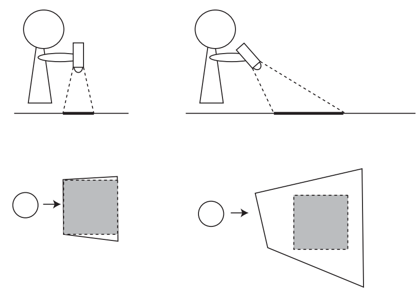

Figure 7: *Size of a source image*

cide on a position, an inclination and a size for the corrected image to be shown inside the projected display. In the CoGAME system, the corrected image is located at the center of the projected display, and the bottom line of the corrected image changes accordingly, to reflect the change in the roll angle of the projector, as shown in Figure 7. The size of the projected image is kept approximately constant, based on the image size when the player projects vertically. This is because the cognitive load on the players would increase if the size of the image changed as the pose of the projector and the distance between the ground plane and the projector changed. To allow a player to resize the images depending on the pose of the projector may make the game more functional, but it may be more complicated to play and robot control may be more difficult. Therefore, the system does not allow players to resize an image, instead keeping the size of the projected image constant.

If the system is to correct the projected image in real time, there is a tradeoff between the accuracy of the correction and the processing time. Hence, the accuracy depends on the computation power of the mobile PC. We decided on a tradeoff based on the accuracy to which the system could correct the projected image in a short constant time (about 0.1 seconds). Because the data from the acceleration sensor includes noise due to hand tremor, the projected image would wobble if the system sensitively corrected it in real time, which would make it difficult for human to interpret. Therefore, the system smoothes the sensor data by averaging its past data.

4.2 Robot configuration and control

A turtle robot has two motors and two wheels that can be separately controlled. The size of the robot is about 10 cm (length) $\times$ 8 cm (width) $\times$ 7 cm (height). It can move at approximately 6 cm/s. For easier recognition of the robot, three infrared LEDs are attached

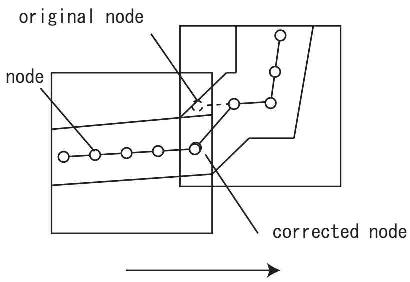

Figure 8: *Smooth connection of projected roads*

to the shell of the turtle robot. Infrared filters are attached to the client cameras. The system can easily identify the robot's position by recognizing the infrared LED. Moreover, by blinking the head LED on and off, the system can also recognize the robot orientation. (Other LEDs are on continuously.)

4.3 Connecting of the roads

A road image can be considered in terms of a list of nodes, with each node being connected to nearby ones, as shown in Figure 8. At any time, the robot should be moving towards one of these nodes. After the system estimates the relative position and orientation of the robot, it determines whether the robot is on the image's road. If the robot is not on the road, the robot moves toward the nearest node. If the robot is on the road, the system moves the robot along the road toward the end opposite to the first position on the road.

When two players connect their roads, the system corrects the road images to achieve a smooth connection. Assume that *player A* is controlling the robot. *Player B* will project an image whose road connects to the end of *player A*'s road, and will then wait for the robot to arrive. In the overlapped region of the two images, the system moves the start node of *player B*'s road to overlay the end node of *player A*'s road. While both players' cameras are capturing the robot, as it approaches the road junction, the system keeps repeating this operation. This procedure enables players to easily connect their roads, without needing manipulations to accurately connect their roads, even when the interface has to deal with wobbles due to hand tremor.

5 Conclusions and Future Work

In this paper, we have proposed the "manipulation by projection" technique. This technique has been applied to allow a user to intuitively control a robot by using a camera and a projector: We have developed an interactive game called CoGAME in which players collaboratively lead a turtle robot to its destination. Through the development of CoGAME, we have implemented a projection correction method, the robot control method, and the method for connecting projected images smoothly.

CoGAME has been demonstrated at "International collegiate Virtual Reality Contest in 2006" held in Tokyo , and the Open Campus event in the University of Tokyo. We have received overall positive feedback from people (ranging from children to the elderly) who tried CoGAME.

In our future work, we will explore possibilities of different applications, such as a system for supporting collaborative tasks, and evaluate the effectiveness of the proposed technique.

6 Acknowledgments

This research has been supported by the 21st Century COE Program "Electronics in Future Generations" from the Japan Society for the Promotion of Science, and by the Core Research for Evolutionary Science and Technology of the Japan Science and Technology Agency. We thank the "International collegiate Virtual Reality Contest (IVRC)" executive committee for their constructive feedback on the CoGAME project.

References

CAO, X., AND BALAKRISHNAN, R. 2006. Interacting with dynamically defined information spaces using a handheld projector and ap pen. In *The annual ACM symposium on User interface software and technology*, 225–234.

MIYAHARA, K., INOUE, H., TSUNESADA, Y., AND SUGIMOTO, M. 2005. Intuitive manipulation techniques for projected displays of mobile devices. In *ACM CHI2005 Extended Abstract*, 1881–1884.

RASKAR, R., VAN BAAR, J., BEARDSLEY, P., WILLWACHER, T., RAO, S., AND FORLINES, C. 2003. ilamps: Geometrically aware and self-configuring projectors. *ACM Transactions on Graphics (TOG) 22*, 3, 809–818.

SHOJI, M., MIURA, K., AND KONNO, A. 2006. U-tsu-shi-o-mi:the virtual humanoid you can reach. In *SIGGRAPH2006 Emerging Technologies*.

SIEMENS. Cell phone with builtin projector. www.physorg.com/news3505.html.

SUGIMOTO, M., KOJIMA, M., NAKAMURA, A., KAGOTANI, G., NII, H., AND INAMI, M. 2005. Augmented coliseum: Display-based computing for augmented reality inspiration computing robot. In *SIGGRAPH 2005 Full Conference DVD-ROM Disk1 Emerging Technologies*.

TOSHIBA, 2006. Tdp-ff1a(j). http://www.toshiba.co.jp/vis/lineup/portable/ff1/index.html.