Development of a City Presentation Method by Linking Viewpoints of a Physical Scale Model and VR

Toshiki Tokuhara¹, Tomohiro Fukuda², Nobuyoshi Yabuki³.
¹,²,³ Graduate School of Engineering, Osaka University, Japan
¹,²,³ (http://y-f-lab.jp/)
¹ tokuhara@it.see.eng.osaka-u.ac.jp, ² fukuda@see.eng.osaka-u.ac.jp, ³ yabuki@see.eng.osaka-u.ac.jp.

Abstract. The design of a city has a great influence on its society. Therefore, current/future cities must be understandable by everyone, regardless of their ability to use technology. Various tools have been used to show urban spaces. The authors focused on SCMODs (physical scale models), and VR (Virtual Reality). These are three-dimensional and intuitive expression methods. In this study, a city presentation method offering a united operating environment linking viewpoint information between a SCMOD and VR is developed and evaluated. Photogrammetry acquires aspect information with a laser pointer, an AR marker and a web camera. To evaluate this method, 36 testees answered a questionnaire after experiencing the method. The testees evaluated the method positively.

Keywords. City presentation; physical scale model; VR; TUI; Photogrammetry.

Introduction

Not only specialists such as architects, city planners, and administrators, etc. but also non-specialists such as local residents, users, and civilians, etc. participate when a city is designed. Moreover, a communications design movement for cities that is called “Civic Pride” has arisen in Barcelona and Amsterdam, etc. further improving the value of these cities and making the citizens feel proud. Therefore, finding presentation methods to understand current/future cities intuitively is becoming increasingly important in all fields.

Up to now, drawings, sketches, CG perspective drawings, SCMODs, and VR have been used as methods of showing a city. SCMODs and VR are three-dimensional expression methods, and enable intuitive understanding. A SCMOD is a three-dimensional object into which a real space is reduced according to a constant ratio. The strengths of a SCMOD are that the user touches the model directly, that several people can examine it at the same time from arbitrary viewpoints, and it allows users to gain an understanding of the entire city. On the other hand, the weaknesses are the limit of expression caused by the reduction, a limitation of the range of production, and that study from the pedestrian viewpoint is difficult. VR is a three-dimensional space expressed in the virtual environment of the computer. The strengths of VR are that a realistic expression by texture mapping is possible, that study from the pedestrian viewpoint is easy, and it can express dynamic urban elements such as people and cars. On the other hand, the weaknesses are that it is impossible to touch directly,
and that the sense of distance is elusive. Both the SC-MOD and VR have strengths and weaknesses respectively as mentioned above. In the current presentation, they are separately used. Might each weakness be rectified by using the two systems together?

In this study, a city presentation method offering a united operating environment linking viewpoint information on a SC-MOD and in VR is developed. The photogrammetry method is adopted as technology to link aspect information. A laser pointer and a web camera are used as input devices.

Previous Studies

In the study and the presentation of city design, the keyboard and the mouse are generally used to move the viewpoint of VR. However, Fukuda et al. (2006) reported that this is an impediment to non-specialists because it is difficult for them to operate VR by the use of the keyboard and the mouse.

The tangible user interface (TUI) has been the object of much research as a possible solution to this problem (Ishii and Brygg, 1997; Rom and Surapong, 2009). Tonn et al. (2008) developed an interface with which users can operate a 3D CAD model on a real scale with a laser pointer and 3D projector. Fujimon et al. (2004) developed a system that displayed VR contents seen from an avatar after having designed a sensor that could acquire the location information as an avatar of the operator. Moreover, Nagakura et al. (2006) developed an interactive space browser for architectural designs. Moving its lightweight LCD panel over the plan of a building drawing displays a 3D interior view of the building. However, it is difficult to apply these systems to a three-dimensional SC-MOD that is the object of this study because the VR interface of these systems targets flat planes such as maps and drawing. Seichter et al. (2004) developed a system to display virtual 3D models using AR-Toolkit and HMD. This system can use SC-MODs and virtual 3D models in an arbitrary mixture. But the system is not able to present a pedestrian’s viewpoint.

Development of a City Presentation System

Figure 1 shows the whole image of the city presentation system developed in this study. First of all, the
user specifies two arbitrary points where a viewpoint and a main object are defined on the SCMOD. The user defines the viewpoint and the main object with the laser pointer by pressing the button on it. Then, a VR image that looks at the main object from the viewpoint defined on the SCMOD on the display is drawn through the laser optical point detection flow, coordinated system conversion flow, and the VR drawing flow. The web camera (two million pixels) is set up 1m above the SCMOD. Next, the laser optical point detection flow, the coordinate system conversion flow, and the VR drawing flow are described.

The laser optical point detection flow
First, the laser optical point detection flow saves an image as a “reference image” from every five images shot by the web camera; the other images are “judgment images”. On the last judgment image, a pixel whose brightness (0-255) is higher than the reference image is made a laser optical point candidate by using the background difference function of OpenCV. If the judgment image has optical point candidates above the “N” threshold which defines the number of pixels with changed brightness, the image is set aside to avoid incorrect detection caused by jiggling of the user’s hand.

The brightness of the pixels indicated by the laser pointer is very low and the brightness of the area surrounding the pixel is also low in contrast. The optical point candidate (the center pixel in Figure 2) is picked up if the absolute value of the difference between the brightness value of the optical point candidate and the brightness value of its 3x3 surrounding pixels (these are slash pattern pixels in Figure 2) is within 20. Next, the brightness of the pixels picked up is deducted from the brightness of each three outside the circuit of 15x15 surrounding pixels (these are gray pixels in Figure 2). If the difference of the brightness value is 30 or more, the candidate is detected as the laser optical point and receives the coordinate value as (LaserX, LaserY).

The coordinate system conversion flow
In the coordinate system conversion flow, the coordinate values (LaserX, LaserY) are converted into coordinate values (ModelX, ModelY) in the coordinate system of the SCMOD, and, in addition, they are converted into coordinate values (VRX, VRY) in the VR coordinate system (VCS). Because the screen coordinate system of the web camera (SCS) and the coordinate system of the SCMOD (MCS) are separately defined, the SCMOD can be freely moved.

The original SCS point is the upper left of the web camera image. To define MCS, ARToolkit and marker (80x80mm) are used. The positional grasp function of the marker in the ARToolkit can obtain

![Figure 2 Pixels searched for detection of optical point (1 block = 1 pixel)](image-url)
the coordinate values of the marker's vertexes and center. The original point of MCS is defined as the center of the marker. Moreover, this function can obtain information on the position and tilt of the marker as a transform matrix (Kato and Billinghurst, 1999). This matrix consists of 3x4 elements. In this matrix, the upper left 2x2 elements show inclination in a perpendicular direction. In this flow, the 2x2 elements of the transform matrix are defined as a rotation matrix. The detailed procedure will be described.

To convert the original point of MCS to the original point of SCS, (LaserX, LaserY) is subtracted from the coordination values of the center of marker (MarkerX, MarkerY), and multiplied by the rotation matrix (Equation 1). The unit of the coordinate values (RotatedX, RotatedY) is the pixel. The distance between the center of the marker and the vertex in the image is measured (Equation 2). This distance corresponds to half the length of the marker's diagonal line (40 \text{mm}). So the conversion coefficient (mmPixRatio) is obtained by dividing (40 \text{mm}) in the result of equation 2 (Equation 3). In addition, the mmPixRatio is multiplied by (RotatedX, RotatedY) (Equation 4). In this way, the coordinate values (LaserX, LaserY) are converted into coordinate values (ModelX, ModelY).

Next, the coordinate values (ModelX, ModelY) are converted into the coordinate values (VRX, VRY) in the VCS. The values (ModelX, ModelY) are divided by the scale of SCMOD and by 1,000 for unit conversion (Equation 5). Then the gap of the original point of SCS and VCS, which is a constant number, is revised. The constant numbers are obtained before running this flow and are saved as (ReviseX, ReviseY). Constant numbers (ReviseX, ReviseY) are added to (ExpandedX, ExpandedY) (Equation 6).

\[
\begin{pmatrix}
R_{11} & R_{12} \\
R_{21} & R_{22}
\end{pmatrix}
\begin{pmatrix}
LaserX - MarkerX \\
LaserY - MarkerY
\end{pmatrix} =
\begin{pmatrix}
RotatedX \\
RotatedY
\end{pmatrix}
\]

\[
\sqrt{(MarkerX - VertexX)^2 + (MarkerY - VertexY)^2} = \text{Distance}
\]

\[
\frac{40\sqrt{2}}{\text{Distance}} = \text{mmPixRatio}
\]

\[
\text{mmPixRatio} \cdot \begin{pmatrix}
RotatedX \\
RotatedY
\end{pmatrix} = \begin{pmatrix}
ModelX \\
ModelY
\end{pmatrix}
\]

\[
\begin{pmatrix}
ModelX \\
ModelY
\end{pmatrix} = \frac{1}{\text{scale} \times 1000}
\begin{pmatrix}
\text{ExpandedX} \\
\text{ExpandedY}
\end{pmatrix}
\]

\[
\begin{pmatrix}
\text{ExpandedX} \\
\text{ExpandedY}
\end{pmatrix} + \begin{pmatrix}
\text{ReviseX} \\
\text{ReviseY}
\end{pmatrix} = \begin{pmatrix}
\text{VRX} \\
\text{VRY}
\end{pmatrix}
\]
In this way, (VRX, VRY) is obtained. In this study, the height value is assumed to be a fixed value (VRZ).

Then the coordinate values of the viewpoint and the main object are output to a file named “Laser Position.dat”. In this file, the coordinate values of the viewpoint and the main object are continuously written (Figure 4). (5)

**The VR drawing flow**

In the VR drawing flow, the VR software “UC-win/Road” (ver.3.4.11) is used. The authors developed plug-in software to put a camera in the VR at the position the user specifies. The plug-in software inputs a file “Laser Position.dat” and memorizes all the numbers described in the file, ten characters per step. These numbers are substituted into the camera information in the VR. Then “Laser Position.dat” is deleted.

**Evaluation of the Developed System and Results**

Accuracy validation and a usability evaluation of the developed system were carried out. The reduced scale of the SCMOD made for the verification was 1/1000, and the area of real scale was a 700m×610m space in the center of Osaka (Figure 5, Figure 6, Figure 7).

First of all, after shining the laser onto the intersection of the grid pattern that set up the marker, the accuracy validation compared the value obtained with the developed software and the value determined with the measure. As a result, an error margin
of about 10mm on average was confirmed (Figure 8, Figure 9). Because the reduced scale of the SCMOD was 1/1000, this corresponded to an error margin of 10m at normal scale.

Next, a usability evaluation was carried out from November to December in 2009. There were 36 testees. According to gender, 23 were male, 4 were female, and 9 did not state their gender. According to age, 18 people were in their 20s, 3 were in their 30s, 4 were in their 40s, 2 were in their 50s, and 9 did not state their age. Specialists and non-specialist were included among the testees. There were 15 persons who had experience of presentations using SCMODs and/or VR, 12 persons with no such experience, and 9 persons gave no information on this point. Figure 10 shows distribution of testees.

After experiencing the system, the testees answered a questionnaire on ease of operation, accuracy, and the response speed. The authors analyzed the data using weighted average (maximum weight is 4) and t-test. Table 1 and Table 2 show the analysis according to the testees' age and experience of presentation. The authors considered testees who experienced a presentation with SCMODs or VR to be specialists. A high appraisal was obtained overall as a result of the analysis. Moreover, it was clear that specialists evaluated this system higher than all other testees in terms of the response speed (significant difference-5%).

Finally, the comments of testees are described. Some testees said this system made it easy to
operate viewpoints in the VR; others said they forgot the viewpoint they had defined and lost track of where they were. This reason was that the point directed by the laser pointer disappears. It is important to develop a function to allow users to understand where the viewpoint is.

### Conclusion

This study succeeded in developing a city presentation system that offered a united operating environment by linking viewpoint information of a SCMOD and VR. The photogrammetry method was adopted to link viewpoint information, and a laser pointer and a web camera were used as input devices. Next, validation of the accuracy and evaluation of the usability of the developed system were carried out. As a result, it was confirmed that an error margin of about 10mm existed in the model coordinate system. Moreover, when the usability was evaluated, specialists evaluated this system higher than all other testees in terms of the response speed.

Future works could investigate how the accuracy of the laser optical point detection flow might be improved, could attempt to improve the interface device, and could consider a system using several web cameras that could be developed for use with larger SCMODs.

### Acknowledgements

A portion of this research was done with the World 16 project and with the assistance of subject number 217604690 of the 2009 Grant-in-Aid for Scientific Research (Japan Society for the Promotion of Science).

### References


