Linked QTVR System for Simulating Citizens’ Strolling Around Activities

Mitsuo Morozumi
(Kumamoto University)

Yuichi Sueshige
(Creative Laboratory, Tokyo University of Technology)

Tadashi Uchiyama
(Kobe Design University)

Susumu Inoue
(Kumamoto University)

Abstract

This paper discusses a city model using the linked QTVR (cylinder-VR) technique to study relationships between people’s behavior while strolling around downtown shopping districts, and visual stimuli in each part of the district. At the beginning of the study, because a large city model was required, a major obstacle was the necessity of taking a large number of photographic images of the area, especially images with wide vertical view angles. The objectives of this paper are to: (a) discuss the major features of the simulation system and its uses for the study, (b) discuss a refined method of producing a cylindrical image of QTVR allowing the operator to look up almost 56 degrees, and (c) evaluate the performance of the developed system through experiment by 30 university students who know the district well.

1 Introduction

An important task of urban design is to provide an attractive environment in which both strangers and frequent visitors to a city may be stimulated into diverse pleasurable activities such as window-shopping and strolling. Past studies on the relationship between citizens’ activities and visual information given by the environment, suggest that such studies can provide general guidance for urban design (Arthur and Passini, 1992). However, it is not at all easy to collect enough data for such studies; observing people’s behavior in a physical environment requires time and energy. It is also difficult to identify the factors that influence behavior.

Limiting subject of the study to observing the relationship between visual stimuli and the behavior of people, the authors started to develop a prototype of an interactive visual simulation system in which one can virtually experience window-shopping and strolling activities in a specific area. After comparing several techniques of visual representation, such as physical models with model scope (Burden, 1985), VRML (Kaga, 1998; Yeung, et al. 1998)\(^1\), linked serial video images (Strong and Woodbury, 1998) and linked QTVR movies (Maver, 1998), the author selected the latter media. (Figure 1)
Figure 1 Image of a city model using linked cylindrical QTVR

Reasons for this choice include:

1) It has an excellent capacity for creating "virtual visits" to city environments (Maver, 1998), yet it works on a common PC system.

2) It is easier to operate than a VRML player, which is an important consideration in a system for simulating strolling around activities where speedy movements are sometimes needed. It allows examinees not only to select their route freely along any course in the city model, reacting to visual stimuli as they go, but also to look up and around and move back and forth along sections of street just by using mouse actions. They can even stare at objects using a zooming utility.

3) Though Strong and Woodbury (1998) pointed out that linked still images tend to lack a sense of sequential movement, the authors decided to test whether linked QTVR movies of a place defined at short intervals could offset this disadvantage.

4) Since a QTVR movie uses photographs of the study area as data, it has high reproducibility of visual information compared to other CG techniques, yet still allows the possibility of graphically rearranging parts of the model to test other design considerations in future studies.

In 1999, the authors built a linked QTVR system for a downtown shopping area of Kumamoto City, Japan (System-I) (Morozumi, 2000). After 10 cases of simulation using strangers to the site, and 10 other cases using people who knew the site very well, it became clear that the system could simulate citizens’ behavior in regard to window-shopping and strolling activities in the area. But several participants also pointed out that the lack of an ability to look upwards gave them an unnatural impression, and that the area of the model was a little too narrow for the simulation. The authors initially planned to build a model for a much wider area, but the time necessary for taking photographs prevented such a plan.

Considering that the application of linked QTVR to large city models would inevitably require the development of an effective method of quickly taking a large number of photographs with a wide vertical view angle, the authors started exploring various photography techniques. The use of wide-angle lenses, including the fish-eye lens, is one of the most effective solutions.

This paper discusses major features of the simulation system and its uses for the study, gives a summary of comparisons of various methods of taking photographs for QTVR models, and finally, makes an evaluation of the performance of the developed system applying the proposed photography techniques (System-II).
2. The simulation system using linked-QTVR

2-1. Major features of the simulation system

The authors took a 500m x 700m central shopping district of Kumamoto City for simulating people's behavior patterns. After preliminary studies, the authors decided to provide a linked QTVR model at every 10-meter point in the streets, in other words, the system needed at least 780 cylindrical models, including sections of wide roads with traffic, which required two rows of linked QTVR models along pedestrian paths located on each side of the road. However, System-I provided only 320 cylindrical models. (Figure 2)

![Figure 2](image-url)

2-2. Method of observing people’s behavior patterns

The authors asked a pair of examinees for each case to manipulate the simulator, sitting in front of a PC, and asked them exchange remarks on what they saw and what they thought, just as they would do if really walking in the street. Making conversation is an effective approach for externalizing thoughts in a natural way (Rogoff, 1991). The authors used a VCR system and tape recorder to record how examinees viewed the city model represented with the linked QTVR model, and what kind of comments they made at each point. (Figure 3)

![Figure 3](image-url)
2-3. Procedure of building city model for the Simulator-I

In building a city model for Simulator-I, the authors used the QTVR Authoring Studio (QTVRAS) to produce a QTVR model from photographs taken in different directions at each point, as well as to set necessary links among models. The QTVRAS normally requires 12 shots at each point, turning a 35mm film camera or the equivalent along the vertical axis on a tripod at 30 degree intervals, using a normal lens with the camera set at the horizontal position.

Some of the problems the authors faced with this method can be summarized as follows:

1) The upper and lower vertical angle of each image is limited to 28 degrees to the horizontal surface. This means that, when one takes a photograph in the center of a 12m wide street, it cannot include the second floor portion of adjacent buildings in the image. Therefore System-I was limited as a simulator because it lacked sufficient ability for looking up.

2) It is essential to be able to rotate the camera along the vertical axis on a tripod. But such camera setting requires time. Even after one gets accustomed to the operation, it normally requires 1 minute for camera setting and 4 minutes to take 12 shots. This means that it requires 5 x 780 = 3900 minutes or 65 hours, at best, to take the necessary photographs of the study area. As a result, the authors had to compromise with 320 QTVR images for System-I.

3) In some sections of street, car traffic did not allow authors even 5 minutes for taking photographs. In such sections, System-I used two linked 180 degree QTVR images instead of one cylindrical image. This arrangement prevented easy operation of the system.

3. Considerations for Technical Improvements

3-1 Alternative methods of taking photographs

The authors explored two separate areas: reducing time required for camera setting, and finding a good combination of lens and camera settings that could reduce the number of shots necessary for a QTVR model while still providing images of wide vertical angle and fine resolution.

For the latter, the authors compared four methods, assuming the use of high-resolution digital cameras with wide angle lenses, including fish-eye lens.

A) Taking four photographs with a fish-eye lens, setting the camera at the horizontal position.

B) Taking two photographs with a fish-eye lens, setting the camera at the horizontal position, and then editing the photos to CubicVR images.

C) Taking ten photographs with a 22mm wide converter lens, setting the camera at the horizontal position.

D) Taking six photographs with a fish-eye lens set at a 200% zoom position, setting the camera at the vertical position.

As the authors could not find commercial software able to convert a series of images taken with a fish-eye lens or a wide converter lens to one panoramic image, they developed the necessary software for A, C, and D, with the collaboration of a software engineer. Theoretically, two fish-eye images could cover the image area in all directions, but it became clear that the QTVR image tended to lose resolution in several parts while converting images composed of parts. For that reason, in method D, the authors set the fish-eye lens at a 200% zoom position to reduce image distortion.

3-2 Comparison of alternatives

Table-1 summarizes the basic data for each alternative. In terms of the vertical angle captured in its images and the number of shots required to produce the QTVR, method B was the best. A and D also showed a good deal of improvement compared to the method used for System-I. However, the QTVR model produced from images A and B could not show enough resolution even though the authors used...
a three-million-pixel digital camera. In both cases, the image resolution was much poorer than that for System-I. One could not identify large displays in a shop located 6 to 8m away, even when one used a zoom up utility. (Figure 4) D provided almost the same resolution as System-I. In terms of resolution of graphic image, C was the best, but it needed 10 photographs. In addition, the file size of each QTVR model would be another barrier when one needed to build a large system. The authors did not select B because there were no tools that could define links among CubicVR models. 

<table>
<thead>
<tr>
<th>System-‡T</th>
<th>Prototype A</th>
<th>Prototype B</th>
<th>Prototype C</th>
<th>Prototype D</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lens</td>
<td>normal</td>
<td>fish-eye</td>
<td>fish-eye</td>
<td>wide</td>
</tr>
<tr>
<td>Number of shots</td>
<td>12</td>
<td>4</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td>Quality of image</td>
<td>good</td>
<td>bad</td>
<td>bad</td>
<td>excellent</td>
</tr>
<tr>
<td>Looking up utility</td>
<td>28°</td>
<td>61°</td>
<td>90°</td>
<td>31°</td>
</tr>
<tr>
<td>File size</td>
<td>0.6MB</td>
<td>1.9MB</td>
<td>1.5MB</td>
<td>24.1MB</td>
</tr>
</tbody>
</table>

Table 1 Evaluation of alternatives

Figure 4 Quality of images of alternatives

Figure 5 Interface of Systems I and II
Comparing this data, the authors selected method D for the new system. It could reduce the necessary shots to half those needed by System-I, while its ability for looking up was twice as much as that of the System-I. (Figure 5, Figure 6)

3-3 Other camera setting devices

In order to reduce time for camera setting, the authors introduced the following devices when taking photographs with method D. (Figure 7)

1) A panorama adjuster, which can be purchased at any camera shop\(^7\), reduced the time taken for camera setting. First it is set on a tripod stand, and its horizontal level is adjusted by manipulating two tripod handles with the help of a spirit level. After this one can easily rotate the camera along the vertical axis, even when the tripod spindle is not vertical, just by rotating the turntable on the adjuster, using the scales attached to it as a guide.

2) The authors produced an L-shaped bracket to attach the digital camera in a vertical position to the turntable of the panorama adjuster.
3) The authors also attached a sunshade plate to the digital camera, because direct sunlight tended to enter the fish-eye lens.

3.4 Procedure for producing linked QTVR movies

1) Since each shot is taken in a vertical position, it is necessary to rotate every photographic image 90 degrees using graphics software such as Adobe Photoshop.

2) Input six images taken at each point to specially developed software, which has the ability to rectify image distortion, to cut out the major portion of the rectified image, and then to edit a series of cut out images in the form of panoramic images that the QTVRAS can read. (Figure 8)

3) Produce a QTVR movie using the Panorama-maker of the QTVRAS

4) Define necessary links among QTVR movies using the Scene-maker QTVRAS to form Simulation System-II.

Figure 8 Interface of the developed program

3.5 Production of Simulation System-II

The authors built System-II, consisting of 779 QTVR movies, in one month, using two sets of digital cameras and one PC for data processing. The average time for camera setting and photo taking at each point was reduced to 2.5 minutes. It took 40 hours to finish all necessary photographs, including travel time in the study area. As the vertical angle of the photographic image was doubled, the user of the system could easily observe the third floor level of buildings standing along a 12m-wide street (Figure 6). System-II also succeeded in eliminating the use of 180-degree QTVR models even at points of heavy traffic because photographic images could be taken in a short time.

4. Evaluation of System-II by users of the system

4-1 Framework of Experiment

The authors conducted two separate sets of experiments to evaluate the applicability of System-II as a simulation system.
1) The authors asked 30 students who knew the area very well to simulate 10-minute virtual walks using the two systems, and then to complete questionnaires rating their impressions of the virtual walks, the quality of visual information, and the operational performance, using 7 SD scales (from -3 to +3) (Table 2). To avoid the effect of order in the system test, students were subdivided into two groups that would test the two systems in a different order. Without asking students for direct comparisons, the authors compared two sets of answers to the same questionnaires, independently asked soon after the relevant test.

2) Authors asked 12 pairs of students who knew the area very well to simulate a 30-minute virtual walk using System-II, after a 30-minute real walk in the study area, and then asked them to answer the same questionnaires used in the first experiment to check how the evaluation changed.

3) The authors also asked students to write a short comment after each answer to amplify the reasons for their responses.

Table 2 List of questions

<table>
<thead>
<tr>
<th>Q1: Did the system give you enough utility for virtually experiencing window-shopping and strolling around activities? (Not enough, or enough)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q2: Did the system give you an atmosphere of the area when you made virtual visits to different parts of the downtown area? (Not enough, or enough)</td>
</tr>
<tr>
<td>Q3: Did you enjoy the virtual walks in the downtown area? (Not at all, or Enjoyed Enough)</td>
</tr>
<tr>
<td>Q4: Did the system give you enough resolution to gain the information necessary to enjoy the virtual walk, such as the size of information on advertisement boards, or window displays in shops, etc? (Not enough, or Enough)</td>
</tr>
<tr>
<td>Q5: Did you easily notice the location of intersections or sharp curves when you entered them while operating the system? (Not easy or Easy)</td>
</tr>
<tr>
<td>Q6: Did the system provide sufficient performance in simulating looking around activities necessary for window-shopping and strolling? (Not enough, or Enough)</td>
</tr>
<tr>
<td>Q7: Did you lose a sense of direction or location while in the simulation? (Seldom, Quite often)</td>
</tr>
<tr>
<td>Q8: What do you feel about the eye level at which the pictures were taken? (Unnatural or natural)</td>
</tr>
<tr>
<td>Q9: How do you feel about the size of the visual area displayed on the screen? (Enough, or Narrow)</td>
</tr>
<tr>
<td>Q10: How do you feel about the operation of the system for turning to the left or right at an intersection? (Unnatural, or Natural)</td>
</tr>
<tr>
<td>Q11: How do you feel about the operation of the system for looking back and then walking back? (Unnatural, or Natural)</td>
</tr>
</tbody>
</table>

4.2 Comparison of two systems

The evaluation of simulator-II exceeded that of simulator-I in questions asking for impressions of the virtual walks (Q1-Q3). Answers to Q1 (utility level of virtual walk) and Q3 (enjoyment level of virtual walk) were especially positive. Reviewing the free comments, it became clear that System-II gained a higher evaluation because of improvement in its looking up and looking around ability. In other words, the increase in the vertical image size, and the elimination of the 180 degree QTVR models. (Figure 9)

In questions that related to the quality of visual information (Q4-Q7), the distribution of scores was large, except those for Q6 (looking around operation). Evaluation seems to have varied, depending on whether the student used zoom up utilities frequently to get details of the image, or whether they just simulated walking around. There were only slight differences between the evaluation of the two systems in Q4 (resolution on images) and Q5, because the systems used images of almost the same resolution. However, it is apparent that the improvement in looking around operation in System-II influenced the score for Q6. According to the free comments, improvements in looking up and looking around
operation of System-II helped the positive shift of scores for Q7 (sense of direction and location) and Q5 (recognition of intersections). (Figure 10)

The average score for Q8 (height of eye level) and Q9 (size of visual area display) was only slightly different between the two systems, but scores for Q10 (turning operation), and Q11 (looking back operation), especially the latter, were much better for System-II because for the same reasons as above. (Figure 11)
evaluations, especially for those questions related to impressions of the virtual walk. Though the authors need to improve the operational performance of the system, especially the resolution of images (Q4), it is possible to claim that System-II has improved its performance as a simulator of virtual window shopping and strolling around activities. (Figure 12)

![Figure 12 Evaluations after the real walks](image)

5 Summary

In this paper the authors have discussed an effective method of taking the photographs necessary to build a city model using linked QTVR movies. After comparing four different methods in terms of number of shots required for each QTVR movie, resolution of images, vertical view angle of images, and size of files, the authors proposed a method of taking six photographs using a high resolution digital camera with a fish-eye lens set at a 200% zoom position, and setting it in a vertical position. This method could reduce field work time as well as allow a better looking up ability for the QTVR system. After building a simulation system (System-II) for Kumamoto city applying the proposed method of taking photographs, the authors tested performance of the system, comparing the virtual walk with the old system (System-I) that has several limits related to the method of taking photographs. They also observed how evaluation changed after a real walk, and concluded that System-II has improved its performance as a system for observing people’s window shopping and strolling around activities.

The authors have already completed eighty 30-minute simulations in a short period, each one corresponding to a real two-hour walk, and are now making further related studies. These include analyzing how a stranger extends his/her activity area, what kinds of visual condition encourage or discourage strangers and frequent visitors to come to the site, and further walks along various sections of street. The authors would like to present the results of these studies in a future paper.

Acknowledgments

The authors would like express special thanks to Mr. Takeshi SAWAI who died several months after completing very effective programs for image conversion used here. Without his dedication, this paper could not have been realized. The authors would also like to thank Mr. Minoru TANAE who worked hard in building the QTVR model, and students of Kumamoto University who participated in the experiment of the prototype system.

References


---

1 Ennis and Maver (2001) developed a comprehensive city model using VRML, after starting the study.

2 In the last paper, the authors tested the prototype I, II, III. The prototype I applied a technique of establishing HTML links among the front, side and rear images of a site, and prototype II and III applied a technique of linked QTVR movies, whose view angle width were respectively 50 degrees and 130 degrees (at the maximum viewing position). For coding simplification, this paper used the name “system I” for the prototype III in the last paper.

3 The author used scan converter (Power Presenter 1280, product of canopus) and video deck (VR-N701, product of FUNAI) for recording image of display to trace how people moved, viewed and what they stared at. For recoding voice, author used Voice Trek (DS-150, product of OLYMPUS), that has an automatic stop utility, activated when examinees stop talking.

4 QTVR Authoring studio 1.0.1J is distributed by Apple Computer Inc.

5 “Nikon CoolPix990”, that can take a 3,340,000-pixel digital image in fine mode.

6 It is possible to define links with HTML on the www, but it cannot realize enough speed when viewing images because it takes time to load models individually.

7 Panorama adjuster, product of MINETTE Special order is required in most cases.

8 $6m \times \tan 56(\text{degree}) + 1.5m = 10.4m$, where 1.5m is height of the camera.

9 The authors set camera at 1.5m above the ground.