SOAR: SENSOR ORIENTED MOBILE AUGMENTED REALITY FOR URBAN LANDSCAPE ASSESSMENT

TOMOHIRO FUKUDA, TIAN ZHANG, AYAKO SHIMIZU, MASAHARU TAGUCHI, LEI SUN AND NOBUYOSHI YABUKI
Osaka University, Suita, Osaka, Japan
fukuda@see.eng.osaka-u.ac.jp

Abstract. This research presents the development of a sensor oriented mobile AR system which realises geometric consistency using GPS, a gyroscope and a video camera which are mounted in a smartphone for urban landscape assessment. A low cost AR system with high flexibility is realised. Consistency of the viewing angle of a video camera and a CG virtual camera, and geometric consistency between a video image and 3DCG are verified. In conclusion, the proposed system was evaluated as feasible and effective.

Keywords. Landscape assessment; augmented reality; mobile device; geometric consistency; GPS.

1. Introduction
In recent years, the need for landscape assessment for public projects, high-rise buildings or high voltage transmission towers has been growing. A landscape assessment means that a project executor hears various opinions, evaluates them, and includes some of them in the project. They help to create a new, good landscape by inclusion in the project execution in each project phase, such as the conceptual phase, planning phase, design phase and maintenance phase. A review meeting of landscape assessment is carried out on a planned construction site in addition to being carried out in a conference room. A landscape is constituted by two or more elements such as artificial objects and natural objects. It is difficult for stakeholders, such as project executors, academic experts, and residents, to imagine concretely such an image that is three-dimensional and does not exist. Therefore, a landscape simulation
method using visualisation systems such as Computer Graphics (CG) and Virtual Reality (VR) has been developed. Traditional perspective drawing, photo montage, and physical models have the problem that the adaptation to change takes time, compared to other digital landscape simulation methods. Moreover, 3DCG perspective drawing or animation has the problem that it is difficult to review immediately from the viewpoint that the reviewer wants at the time of a meeting. In order to check whether some portions of high structures are visible or not behind the locations of interest from multiple viewpoints in the landscape assessment process, it is necessary to make a 3D model representing the geography, existing structures and natural objects using 3D CAD or VR software. However, it usually requires much time and expense to make such a 3D model. Moreover, since consistency with real space is not achieved when using VR on a planned construction site, it has the problem that a reviewer cannot get an immersive experience.

Thus, in this research, the authors propose a new method using Augmented Reality (AR) which can superimpose an actual landscape acquired with a video camera and 3DCG. If AR is used, a landscape assessment object will be included in the present surroundings, and it will become possible to carry out a landscape preservation study. Thereby, a drastic reduction of the time and expense involved in carrying out 3DCG modelling of the present surroundings can be expected. In AR, realisation of geometric consistency with a video image of an actual landscape and CG is an important feature (Viet et al. 2009, Charles 2010, Ming 2010, Suyang 2010). The following methods are proposed: 1) use of physical sensors such as the Global Positioning System (GPS) and gyroscope, 2) use of an artificial marker, and 3) the extraction of characteristic points. In method 1), in order to realise highly precise geometric consistency, special hardware which is expensive is required (Fukuda et al. 2006, Leon et al. 2009, Watanabe 2011). That is, a problem remains in utilisation since equipment is required that not all people may have. In method 2), geometric consistency is realised using an artificial marker which can be made at low-cost (Yabuki et al. 2011). However, since an artificial marker needs to be always visible by the AR camera, the movable span of a user is limited. Moreover, in order to realise high precision, it is necessary to use a large artificial marker.

In this research, a smartphone that is widely available on the market level is adopted as an AR device. SOAR (Sensor Oriented Mobile AR) system which realises geometric consistency using GPS, a gyroscope and a video camera which are mounted in a smartphone is developed. A low cost AR system with high flexibility is realisable through this research.
2. System development

2.1. DEVELOPMENT ENVIRONMENT OF A SYSTEM

The authors use a standard spec smartphone, a SoftBank 003SH running Android 2.2. Development languages are OpenGL-ES ver.2.0 and Java ver.1.6. The development environment is Eclipse Galileo ver.3.5.

2.2. SYSTEM FLOW

The flow of the developed system is shown below (Figure 1).

1. Rendering of a video image is described. While the CG model realises ideal rendering by the perspective drawing method, rendering of a video camera produces distortion. Therefore, it is necessary to calibrate the video camera using Android NDK-OpenCV.
2. The definition of a 3DCG model which is the target of landscape assessment is described. Firstly, in a 3DCG model, the geometry (.obj) and the unit of the 3DCG model are described. Secondly, in a 3DCG model allocation file, the name, the file name, the position data (longitude, latitude, ellipsoidal height), the degree of rotation angle, and the zone number of the rectangular plane of the 3DCG model are described. Finally, in a 3DCG model the allocation list file, the number of the 3DCG model allocation information file and each name in the 3DCG model arrangement information file are described.
3. A system user selects a 3DCG model that will create a landscape assessment object using the GUI of a smart phone.
4. After the 3DCG model is selected, acquisition of the position information (latitude, longitude, ellipsoidal height) and angle information (yaw, pitch, roll) of the user’s present location is started with GPS and the gyroscope in the smartphone. Moreover, acquisition of a live video image of the present location is started with a video camera.
5. The definition of the position and orientation data on a CG virtual camera which renders 3DCG is described. The position data (longitude, latitude) acquired by the GPS mounted in the smartphone is converted into the coordinates (x, y) of a rectangular plane. Orthometric height is created by subtracting the geoid height (Geospatial Information Authority of Japan 2011) from the ellipsoidal height acquired by the GPS. Angle values of yaw, pitch, and roll are acquired by the gyroscope mounted in the smartphone. A yaw value points out magnetic north. In an AR system, in order to use a true north value, a magnetic declination is acquired and corrected.
6. A 3DCG model is made to superimpose on the live video image of a video camera.
3. Verification of the system

Verification of the developed system is described. In this research, consistency of the viewing angle of a video camera and CG virtual camera and accuracy of the geometric consistency of the video image and 3DCG are verified.

3.1. Consistency with the Viewing Angle of a Video Camera and CG Virtual Camera

In order to realise geometric consistency between a 3DCG model and a live video image, firstly, it is necessary to uniform the viewing angle of CG virtual camera and a video camera. Although the viewing angle of the video camera is a fixed value, since it is generally not shown to the public, the consistent viewing angle is discovered through an experiment.

3.1.1. Experimental methodology

First, both a rectangular parallelepiped physical model and a 3DCG model of the same size were prepared. Next, the physical model was installed at a position where the distance from the video camera would be 1 m, 3 m, and 5 m, respectively in that order. In an analogous way, the 3DCG model was installed in a position where the distance from the CG virtual camera would be 1 m, 3 m, and 5 m, respectively in order. Then the live video image of the physical model and the 3DCG model are rendered simultaneously. The width and height of the physical model and the 3DCG model which are rendered were measured. The viewing angle which showed the smallest value of both differences in dimension was the viewing angle to determine. In this experi-
3.1.2. Results

The size of the 3DCG model was about 0.02 times as large as the physical model when the horizontal viewing angle of the CG virtual camera was 37.1 degrees. The size of the 3DCG model was about 0.0003 times as large as the physical model when the angle was 37.2 degrees. The size of the 3DCG model was about 0.003 times as small as the model when the angle was 37.3 degrees. As a result of the experiment, the horizontal viewing angle of the CG virtual camera obtains 37.2 degrees.

An experimental photograph for which the horizontal viewing angle of the CG virtual camera was set at 37.2 degrees is shown in Figure 2. The ratio (3DCG model / physical model) of the width and height is shown in Figure 3.

![Figure 2. Experiment photograph (Black rectangle: 3DCG; White rectangle: physical model).](image)

![Figure 3. Ratio (3DCG model/physical model) of the width (left) and height (right).](image)

3.2. ACCURACY OF GEOMETRIC CONSISTENCY WITH A VIDEO IMAGE AND 3DCG

3.2.1. Experimental methodology

The parameters for realising geometric consistency are the position information (latitude, longitude, ellipsoidal height) acquired by GPS, and angle infor-
mation (yaw, pitch, roll) acquired with a gyroscope. The accuracy of geometric consistency is determined by combining the residual error of these parameters. First, what kind of detection characteristic each parameter showed was identified. Therefore, a landscape viewpoint place with known position information and angle information was set up. In one experiment, only one parameter was acquired from a device and the remaining parameters set up a known value as a fixed value. A 3DCG model created from the design drawing was made of an existing building to superimpose on the live video image, and the residual error was measured.

Consequently, eight experiments were conducted in this research (Table 1). In experiment No. 1, all the data on the latitude, longitude, ellipsoidal height, yaw, pitch, and roll of the CG virtual camera used known static values. In experiment No. 2, all the data on the latitude, longitude, ellipsoidal height, yaw, pitch, and roll of the CG virtual camera used dynamic values acquired from GPS and a gyroscope slope. The dynamic values acquired from GPS and a gyroscope slope were used only for one parameter among the latitude, longitude, ellipsoidal height, yaw, pitch, and roll of the CG virtual camera, in experiment No. 3 to No. 8; the remainder used the static values.

In this research, the GSE Common East Building of Osaka University (W29.6 m, D29.0 m, H67.0 m) was used as a landscape assessment object. As a landscape assessment viewpoint place, No.14-563 reference point ((latitude, longitude, orthometric height) = (34.82145699, 135.519612, 53.1)) in Suita city, from where the GSE Common East Building can be seen, was used. The distance from the reference point to the centre of the GSE Common East Building was 203 m. An AR system was installed with a tripod at a level height 1.5 m from the ground. An experiment photograph and measuring points of residual error are shown in Figure 4. The experiment condition is shown in Table 2.

![Figure 4. Experiment plan (left), photograph and measuring points of residual error (right).](image-url)
TABLE 1. Parameter setting in the experiment (S: Static value; D: Dynamic value).

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Position information of CG virtual camera</th>
<th>Angle information of CG virtual camera</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>latitude</td>
<td>longitude</td>
</tr>
<tr>
<td>No.1</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>No.2</td>
<td>D</td>
<td>D</td>
</tr>
<tr>
<td>No.3</td>
<td>D</td>
<td>S</td>
</tr>
<tr>
<td>No.4</td>
<td>S</td>
<td>D</td>
</tr>
<tr>
<td>No.5</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>No.6</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>No.7</td>
<td>S</td>
<td>S</td>
</tr>
<tr>
<td>No.8</td>
<td>S</td>
<td>S</td>
</tr>
</tbody>
</table>

TABLE 2. Experiment condition.

<table>
<thead>
<tr>
<th>Experiment time</th>
<th>Sep. 18, 2011, at 12:00 – 15:00</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weather</td>
<td>Fair, occasionally cloudy</td>
</tr>
<tr>
<td>Equipment</td>
<td>Smartphone (003SH), tripod, measure</td>
</tr>
</tbody>
</table>

3.2.2. Calculation procedure of residual error

The calculation procedure of residual error at the measuring point is shown.

1. The AR image was captured, and four corners of the building on the live video image were set to the measuring points of residual error A, B, C, and D. Each difference between the horizontal direction and vertical direction of those points and the 3DCG model was measured in terms of the number of pixels ($\Delta x$, $\Delta y$). This is called a pixel error.

2. From the acquired value ($\Delta x$, $\Delta y$), each difference in the horizontal direction and vertical direction was computed as a meter unit by the formula 1 and the formula 2 ($\Delta X$, $\Delta Y$). This is called a distance error.

\[
\Delta X = \Delta x \cdot \frac{W}{x} \quad (1)
\]

\[
\Delta Y = \Delta y \cdot \frac{H}{y} \quad (2)
\]

$W$: Actual width of an object (m), $H$: Actual height of an object (m), $x$: Width of 3DCG model on AR image (px), $y$: Height of 3DCG model on AR image (px)

3.2.3. Results

- In experiment No. 1, the pixel error was 1.5 pixels or less, and the mean distance error was less than 0.15 m. From this result, the accuracy of the AR
A drawing of a system was found to be high and it is suggested that an object 200 m away can be fully evaluated when a known static value is used (Figure 5 left and Figure 6).

- In experiment No. 2, the pixel error was less than 20 pixels in the horizontal direction and less than 55 pixels in the vertical direction. The mean distance error was 2.3 m in the horizontal direction and 6.3 m in the vertical direction (Figure 5 right and Figure 6).

- In experiment No. 3, only the latitude data of the CG virtual camera used the dynamic value acquired from GPS. The pixel error was less than 23 pixels in the horizontal direction and less than 14 pixels in the vertical direction. The mean distance error was 2.6 m in the horizontal direction and 1.6 m in the vertical direction. In this experiment, the object was north-northeastward from the viewpoint. Therefore, when the latitude was made into a dynamic value, a 3DCG model was drawn so that it might be expanded and reduced (Figure 6).

- In experiment No. 4, only the longitude data of the CG virtual camera used the dynamic value acquired from GPS. The pixel error was less than 97 pixels in the horizontal direction and less than 11 pixels in the vertical direction. The mean distance error was 11.2 m in the horizontal direction and 1.2 m in the vertical direction, and shifted horizontally (Figure 6).

- In experiment No. 5 and No. 7, the vertical pixel error was large. That is, if an ellipsoidal height or pitch is made into a dynamic value, the position of 3DCG model will move up and down (Figure 7).

- In experiment No. 6, only the yaw data of the CG virtual camera used the dynamic value acquired from the gyroscope. The pixel error was less than 115 pixels in the horizontal direction and less than 5 pixels in the vertical direction. The mean distance error was 13.2 m in the horizontal direction and 0.6 m in the vertical direction, and shifted horizontally greatly (Figure 7).

- In experiment No. 8, only the roll data of the CG virtual camera used the dynamic value acquired from the gyroscope. The pixel error was less than 14 pixels in the horizontal direction and less than 5 pixels in the vertical direction. The mean distance error was 1.6 m in the horizontal direction and 0.5 m in the vertical direction, and this parameter had a small influence on the residual error (Figure 7).

Figure 5. AR screen shot of experiment No.1 (left) and No.2 (right).
The mean distance error acquired in each experiment is shown by the absolute value in Figures 6 and 7. Although No. 2 is all dynamic inputs, the X residual error of No. 2 is smaller than the one of No. 6, for example. Namely, it is the result of offsetting the X residual error of No. 2, the X residual error of No. 6 etc. which is + angle, and the X residual error of No. 4 etc. which is – angle.

4. Conclusion

The contributions of this research are as follows:

- The developed AR system has geometric consistency using GPS and the gyroscope with which the smart phone is equipped. Therefore, a user can use it easily and we can describe it as a system with high flexibility.
- The horizontal viewing angle of the CG virtual camera for realising highly precise geometric consistency was 37.2 degrees. Moreover, when a known value was used for the position information, and the angle information on the CG virtual camera as a static value and the distance between the AR system and an object was 200 m, the pixel error was less than 1.5 pixels, and the mean distance error was less than 0.15 m. This value is the tolerance level of landscape assessment. On the other hand, when the dynamic value acquired with GPS and the gyroscope was used, the pixel error was less than 20 pixels in the
horizontal direction and 55 pixels in the vertical direction. The mean distance error was 2.3 m in the horizontal direction and 6.3 m in the vertical direction.

- The characteristic of the residual error which each parameter acquired with GPS and the gyroscope has been analysed. When the error of GPS is large, it is corrected by carrying out direct entry of the numerical value.

A future work should attempt to reduce the residual error included in the dynamic value acquired with GPS and the gyroscope. Moreover, it is also necessary to verify accuracy of the residual error to objects further than 200 m away and usability.

References


