Expert Systems With Applications 55 (2016) 48–58

Contents lists available at ScienceDirect

Expert Systems With Applications

journal homepage: www.elsevier.com/locate/eswa

A 3D GIS-based interactive registration mechanism for outdoor augmented reality system

Wei Huang[a], Min Sun[b,*], Songnian Li[a]

[a] Department of Civil Engineering, Ryerson University, Toronto, Canada
[b] Institute of RS & GIS, Peking University, Beijing, China

ARTICLE INFO

Keywords:
Augmented reality
3D GIS
ARGIS
Registration

ABSTRACT

Registering virtual objects with reference to their corresponding real-world objects plays a key role in augmented reality (AR) system. Although there have been a lot of work on using vision-based method to perform registration for indoor AR system, it is very difficult to apply such registration method for outdoor AR system due to the inability to modify the objects in outdoor environment and the huge range of working area. 3D Geographic Information System (GIS) is capable of providing an outdoor virtual geographic environment where users are located at, which may provide users with a corresponding virtual object for the one in the physical world. In this study, a 3D GIS-based registration mechanism is proposed for outdoor AR system. Specifically, an easy-use interactive method for precise registration was developed to improve the performance of the registration. To implement the registration mechanism, an outdoor AR system built upon 3D GIS was developed, named Augmented Reality Geographical Information System (ARGIS). ARGIS has the capability of performing precise registration in outdoor environment without using traditional vision tracking method, which thus enables users to arbitrarily manipulate the system. A prototype was developed to conduct experiment on the campus of Peking University, Beijing, China to test the proposed registration mechanism. The experiment shows that the developed registration mechanism is feasible and efficient in the outdoor environment. The ARGIS is expected to enrich the applications of outdoor AR system, including but not limited to underground facility mapping, emergency rescue and urban planning.

1. Introduction

Augmented Reality (AR) combines and aligns real-world and virtual objects in a physical environment to serve a wide array of applications, including medical procedures, manufacturing, entertainment, construction and education (Azuma et al., 2001, 1997; Lee, Huang, Huang, Hsieh, & Lee, 2012b; Seo & Lee, 2013). In accordance to the tracking method used, AR system can be classified into indoor AR system and outdoor AR system (Azuma, Hoff, Neely III, & Sarfaty, 1999). An indoor environment is being treated as a finite space that can be registered with specific fixed artificial references (e.g., fiducial markers). Thus, common vision tracking methods can be directly applied to aid in virtual object registration within indoor environment. On the other hand, an outdoor AR system requires Global Positioning System (GPS) and Inertial Measurement Unit (IMU) sensors to aid in the registration process due to the inability to modify the objects in outdoor environment and the huge range of working environment (Behringer, Park, & Sundareswaran, 2002; Feiner, MacIntyre, Höllerer, & Webster, 1997; Piekarski & Thomas, 2002; Thomas et al., 2000). Despite that, the performance of real-time registration can hardly satisfy the demand of object tracking by using GPS and IMU only in outdoor environment because of the systematic error produced by the equipment. Therefore, it is necessary to develop a new method to fill this gap, which enables the outdoor AR registration to be accomplished efficiently and precisely.

Conceptually, the augmented objects in any AR system should have their corresponding objects to be associated in the real world. Thus, 3D Geographic Information System (3D GIS) is able to provide a good digital version of any outdoor environment where users are located at. Traditional 3D GIS emphasizes high quality photorealistic 3D objects (static objects are mainly considered), while such a requirement is not necessary in AR system because users are already situated at the outdoor scene. Generation of 3D GIS models and their corresponding data management process are thereby simplified dramatically. Outdoor AR system only requires

* Corresponding author. Tel.: +86 1062764484.
E-mail addresses: wei1.huang@ryerson.ca (W. Huang), sunmin@pku.edu.cn (M. Sun), snli@ryerson.ca (S. Li).

http://dx.doi.org/10.1016/j.eswa.2016.01.037
0957-4174/

the 3D wireframe models of augmented objects, and the system should place more emphasis on invisible information of physical environment, such as objects size, attribute information (e.g., building names and street names), directions, and geometric characteristics (e.g., distance, area and volume).

In this study, 3D GIS is introduced to AR system to improve the efficiency and accuracy of registration in outdoor environment. The combined system is named as ARGIS. First, a coarse registration process can be achieved by using virtual 3D GIS objects and positioning data collected by GPS and IMU sensors. Then, with the aid of a handheld interactive device, users can perform precise 3D registration of outdoor scenes based on the principle of 3D coordinate transformation. Such a human–computer interaction can facilitate the refinement of registration from passive mode into active operation, resulting in a better flexibility and higher registration accuracy. Moreover, the interactive device also provides a tool which enables users to communicate with each other (e.g., input certain simple commands to request the augmented views of other users).

The paper is organized as follows. Related work is presented in Section 2. Section 3 introduces the architecture of ARGIS. A new interactive method for precise registration is explained in Section 4. The experiment is discussed in Section 5. Conclusion and future work are presented at the end of the paper.

2. Related work

Existing studies, such as Zlatanova and van den Heuvel (2002), Ghadirian and Bishop (2002), Liarokapis et al. (2005) and Uchiyama, Saito, Servières, and Moreau (2009), used 3D GIS to provide auxiliary information of outdoor scenes (e.g., building attributes, road maps and environmental-related information) to the AR system. Another stream of research focuses on the development of registration methods based on the information provided by 3D GIS. These studies, including Thomas et al. (2000), Behringer et al. (2002), King, Piekarski, and Thomas (2005), and Schall et al. (2009a), attempted to track apparent references or objects provided by 3D GIS, and subsequently registered these objects into video streams. Due to the complex nature of urban morphology, it is hard to locate the reference objects and this kind of registration methods may not be practical in outdoor environment.

On the other hand, Jiang, Neumann, and You (2004) and Satoh, Anabuki, Yamamoto, and Tamura (2001) proposed to combine the use of GPS and IMU sensors together with vision tracking, leading to a viable solution for outdoor AR registration. However, the registration does not work well under rapid viewpoint displacement and rotation (Yuan, Ong, & Nee, 2006). Fong, Ong, and Nee (2009) developed a hybrid tracking method that used data collected by GPS and IMU to locate users' position, and subsequently performed registration by tracking pre-defined references. Recently, there have been several attempts using virtual 3D models to improve registration accuracy (Behzadan & Kamat, 2007; Schall et al., 2009b), or GPS and gyroscope signals to stabilize registration process (Azuma, Neely, Daily, & Leonard, 2006; Piekarski & Smith, 2006; Reitmayr, Eade, & Drummond, 2005). However, these methods still require the use of vision tracking, which has several drawbacks: (1) objects with complex geometry (e.g., trees), or situated under non-uniform illumination (i.e., under/over-exposure) are hard to be recognized; (2) vision tracking can only handle static scenes without significant fluctuation, and thus the operation is sensitive to the user's stability; and (3) vision tracking demands powerful computing capability that imposes significant hardware burden for outdoor AR.

Actually, some interactive operations were used to tackle these problems involved in vision tracking. Piekarski and Smith (2006) proposed an interactive method for an outdoor AR system that registered 3D objects in outdoor environment using special robust gloves. However, the proposed system was developed based on tracking the gloves to add 3D objects without any geographic information provided. Azuma et al. (2006) introduced mobile beacons to act as a reference for tracking, and combined the positioning data collected by GPS to register virtual 3D objects. The study only included a simulation experiment without any performance testing in outdoor environment. Moreover, most of the recent researches on mobile AR system (Carmigniani et al., 2011; Vincent, Nigay, & Kurata, 2013) are heavily dependent on the use of predefined artificial signs/targets, which cannot satisfy the registration process in complex outdoor environments.

Recently, some vision tracking independent outdoor AR applications have been developed. For instance, Lee, Dunser, Nassani, Billinghurst (2013); Lee, Dunser, Kim, and Billinghurst (2012a) developed a mobile outdoor AR application for city visualization and a virtual tour to Antarctica AR application. However, most of these applications mainly focus on the variety of related information provided to overlay, without reporting on registration performance. Moreover, they placed efforts on adding some related information around a location or about a building just into the interface of the applications rather than precisely registering them into the corresponding objects in the real world (Li, Weng, Zhou, Hao, & Zhao, 2013; Yovcheva, Buhalis, & Gatzidis, 2012).

In summary, the aforementioned researches either still need to use vision tracking and rely on certain predefined artificial signs/targets to accomplish the registration process or do not perform precise registration. To tackle these issues for outdoor AR, this work aims to develop an outdoor system (i.e., ARGIS) to demonstrate a new registration mechanism without using vision tracking which enables precision registration to be accomplished efficiently in outdoor environments.

3. An overview of ARGIS system architecture

The system architecture of the proposed ARGIS follows a typical 3-tier client/server architecture, which consists of multiple mobile clients, a centralized server and a 3D GIS database, as illustrated in Fig. 1. Each mobile terminal (client) requests data from the server to augment the scene of outdoor environment through wireless network. The server at the back-end searches and fetches data from the 3D GIS database according to the client's location and orientation measured by GPS and IMU and pushes the requested data to the client. The 3D GIS database stores 3D objects (e.g., wireframe model of buildings) so that they can be streamed to the client for precise registration. Furthermore, ARGIS allows each client to share the current augmented images with other clients and the server, and thus, for example, rescuers who work as a team could conduct more accurate rescue operations if they know certain details about other teammates' working scene.

Fig. 2 illustrates how the ARGIS interactive registration works. First, the mobile client system utilizes real time GPS and IMU data to approximately match the virtual objects to the corresponding real-world objects in the physical environment. Then, users can fine-tune the position and/or orientation through the handheld interactive device, which improves the matching accuracy between the wireframe models and the corresponding real-world objects. This process is considered as the precise registration between virtual environment and physical environment. With the introduction of 3D GIS, such precise registration process can be performed by interactively matching the objects in the physical environment (e.g., a house) to the corresponding objects in 3D GIS (e.g., the 3D wireframe model of a house). Comparing with the visual tracking method, this method avoids the video streaming process and thus dramatically reduces the computational consumption of the system and breaks the limitation of work space.

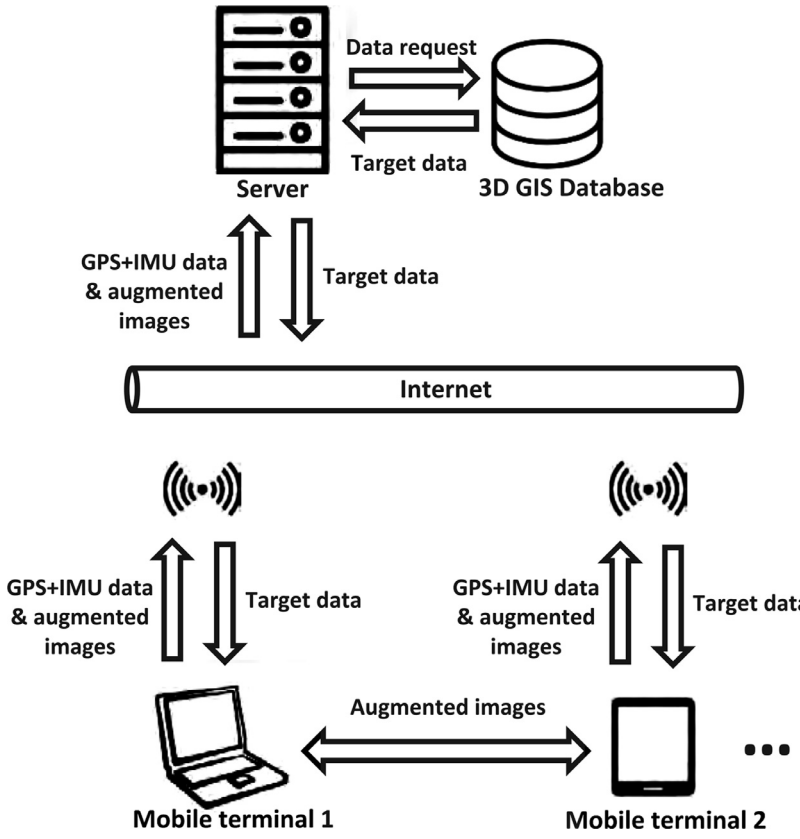

Fig. 1. Overall architecture and design of ARGIS.

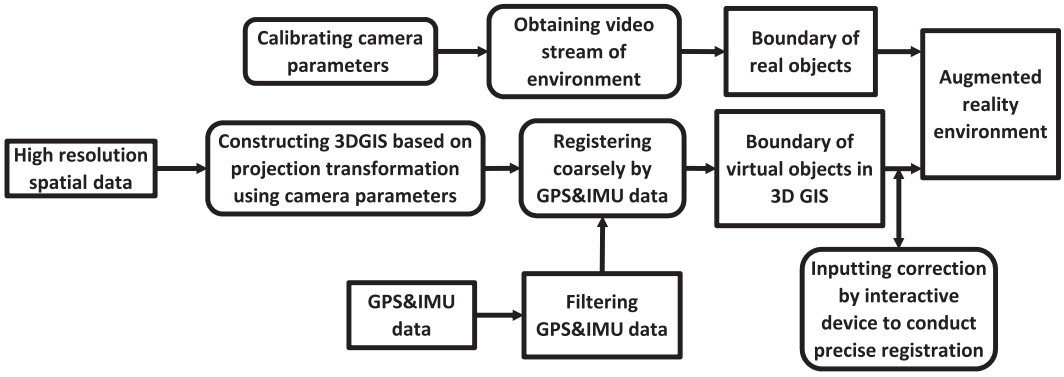

Fig. 2. Methods of ARGIS interactive registration.

When conducting the coarse registration using the data collected by GPS and IMU, the virtual 3D objects may not be well aligned with the real-world objects. As illustrated in Fig. 3, the virtual wireframe model (as highlighted in yellow solid line) appears to be tilted and shifted with reference to the real-world building object (as highlighted in yellow dash line) in the video stream. Such registration error could be caused by two major factors: (1) the geo-positioning accuracy of virtual 3D wireframe model and digital elevation model (DEM), and (2) the measurement error due to the consumer-grade GPS receivers and gyroscopes within dense high-rise urban environment. Though incorporation of wireless cellular positioning technique is a common trend in wearable computing, the geo-positioning accuracy can vary from 10 to 100 m (Mok, Shea, & Yan, 2004; Pun-Cheng, Mok, Shea, & Yan, 2007), which may not fulfill the requirement of precise registration. Since ARGIS is aimed to be designed for use in urban areas, the error from (1) can be removed by using highly accurate 3D models

Fig. 3. Results of coarse registration for the building wireframe model based on GPS and IMU sensors. (For interpretation of the references to color in this figure text, the reader is referred to the web version of this article.)

which can be extracted from other datasets, such as LiDAR (Light Detection and Ranging) data. In this case, the following discussion will focus on how to handle the errors measured by GPS and gyroscope.

4. A new interactive method for precise registration

The use of interactive device plays an important role in ARGIS. Field users can switch between different functions and/or search for different information using the interactive devices and users at the terminal end can communicate with server-side users through the interactive devices. With regards to the interaction, two assumptions have to be fulfilled in advance: (1) if there is no linear relationship between the input correction and output registration compensation, users are not able to handle the manipulation of input correction; and (2) since there are six input corrections affecting the results of registration while only two output compensations (i.e., vertical and horizontal compensations on augmented images) are required, the relationship between input and output parameters has to be established in order to design a set of practical interactive strategies.

4.1. Relationship between input correction and output compensation

After coarse registration, the virtual viewpoint in 3D GIS should be somehow matched to the corresponding real-world viewpoint, and the coordinates of the target point are transformed to the real-world coordinates based on where the user located at (the point of observation). A 3D coordinate system is established in accordance to the point of observation which is set as origin and the horizontal direction ahead of the observer is set as Z (Fig. 4(a)). Assuming there is a target point O in the 3D space (shown in Fig. 4(b)), if the viewpoint of the observer without measurement error is located at C_1, the projection of O on the image should be identified as m. By taking into account the orientation and positioning errors, the measured viewpoint can be referred to C_2 so that the actual projection of O on the image is m'. Accordingly, the projection error can be defined as $dm = m - m'$. Suppose dR and dt are orientation error and positioning error, respectively, i.e., C_2 has both rotation (dR) and translation (dt) from C_1; and also suppose the homogeneous coordinates of point O is $[M1]$, when the observer is stable and the projection parameters are ignored, the relationship between dm and dR and dm and dt is able to be inferred. According to the principle of projective geometry, if C_1 is set as reference origin, the projection point m' caused by errors can be approximately defined as: $m' = f[dR|dt][M1]^T$, where f indicates focal length.

Since dR refers to a small-angle error, it can be represented as radian in terms of rotation amount $[\omega\, k\, \varphi]$ corresponding to the three axes sequenced as $[ZYX]$, and dt can be represented by the vector $[dt_x\, dt_y\, dt_z]$, then

$$m'_x = f\frac{X - \omega Y + kZ + dt_x}{\varphi Y - kX + Z + dt_z} \quad m'_y = f\frac{\omega X + Y - \varphi Z + dt_y}{\varphi Y - kX + Z + dt_z} \tag{1}$$

There exist seven transformation parameters between the virtual and real-world scenes. Scale factor can be assumed as constant, thus it can be ignored. In this case, there only exist six input parameters left being taken into account for manual manipulation, including three positioning parameters, $[X_p, Y_p, Z_p]$, and three orientation parameters, $[e_\omega, e_k, e_\varphi]$. Based on Eq. (1), the relationship between every input parameter and the registration compensation obtained from projection images, represented by dm'_x and dm'_y, can be inferred. Suppose the respective input corrections of rotation angle along Z, Y and X axes are e_ω, e_k and e_φ, the corresponding registration compensations can be obtained as follows:

$$dm'_x = f\frac{-e_\omega Y}{Z} \quad dm'_y = f\frac{e_\omega X}{Z} \tag{2}$$

$$dm'_x = f\frac{(Z^2 + X^2)e_k}{Z(Z - e_k X)} \approx fe_k \quad dm'_y = f\frac{-XYe_k}{Z(Z - e_k X)} \approx 0 \tag{3}$$

$$dm'_x = f\frac{-XYe_\varphi}{Z(Z + e_\varphi Y)} \approx 0 \quad dm'_y = f\frac{-(Z^2 + Y^2)e_\varphi}{Z(Z + e_\varphi Y)} \approx -fe_k \tag{4}$$

Normally, the value of Z of the target point in front of the observer is significantly greater than the value of X and Y, therefore dm'_y from Eq. (3) and dm'_x from Eq. (4) can be ignored comparing with other compensations. For instance, since the reference point for correction is supposed to be selected along the direction of observation, the value of X and Y should be relatively small. Suppose $X = Y = 10$ m, $e_\omega = e_k = e_\varphi = 0.2^\circ = 0.2 \times \pi/180^\circ \approx 0.0035$, $f = 1000p$ and $Z = 50$ m, then, $dm'_x = -0.7p$, $dm'_y = 0.7p$ in Eq. (2); $dm'_x \approx 3.5p$ in Eq. (3); $dm'_y \approx -3.5p$ in Eq. (4).

Suppose the input corrections of position along Z, Y and X axes are X_p, Y_p and Z_p, respectively, the corresponding registration compensations can be derived as follows:

$$dm'_x = f\frac{X_p}{Z} \quad dm'_y = 0 \tag{5}$$

$$dm'_x = 0 \quad dm'_y = f\frac{Y_p}{Z} \tag{6}$$

$$dm'_x = f\frac{XZ_p}{Z^2 + ZZ_p} \approx 0 \quad dm'_y = f\frac{YZ_p}{Z^2 + ZZ_p} \approx 0 \tag{7}$$

Similarly, considering the value of Z is significantly greater than the value of X and Y, the results of Eq. (7) can be ignored comparing with Eqs. (5) and (6). For example, suppose $X_p = Y_p = 0.05$ m, $Z = 50$ m and $f = 1000p$, then $dm'_x = 1p$ in Eq. (5) and $dm'_y = 1p$ in Eq. (6).

From the above equations, it is obvious that there exists a linear relationship between the input corrections and output registration compensations, which leads to the viable mechanism of interactive registration.

4.2. Input corrections

Since the deviation caused by coarse registration can be only visually recognized, the positioning and orientation corrections cannot be precisely inputted at once. Instead, trace values need to be continually inputted to gradually reach the corrections. Therefore, a proper trace value is vital since a large trace value affects the result of precise registration, while a small trace value increases the processing time of correction. To tackle this problem, a velocity function, $H(t)$, is introduced based on selecting a proper incremental input which is obtained through manual testing and prior knowledge.

The output correction is supposed to be d_i, the linear relationship from Eqs. (2) to (6) can be simplified as: $d_1 = k_1 e$ and $d_2 = k_2 p$, where e and p refer to the input correction of rotation and position, respectively; the value of k_1 and k_2 can be computed by Eqs. (2)–(4) and Eqs. (5)–(7), respectively. Angular and positioning incremental input are supposed to be Δe and Δp, respectively, then the relationship can be expressed as

$$d_1 = \sum H_1(t) k_1 \Delta e \quad d_2 = \sum H_2(t) k_2 \Delta p \tag{8}$$

where $H_1(t) = \lambda_1/t$, $H_2(t) = \lambda_2/t$, λ_1 and λ_2 are the two prior values which can be modified during the registration process, and thus ensure the interactive registration to be conducted precisely. By using this velocity function, users are able to efficiently and precisely align the objects without changing the trace value even when the error after coarse registration is relatively great. In other words, the trace value can be automatically decreased by the velocity function (e.g., $H_1(t) = 1/t$) over time.

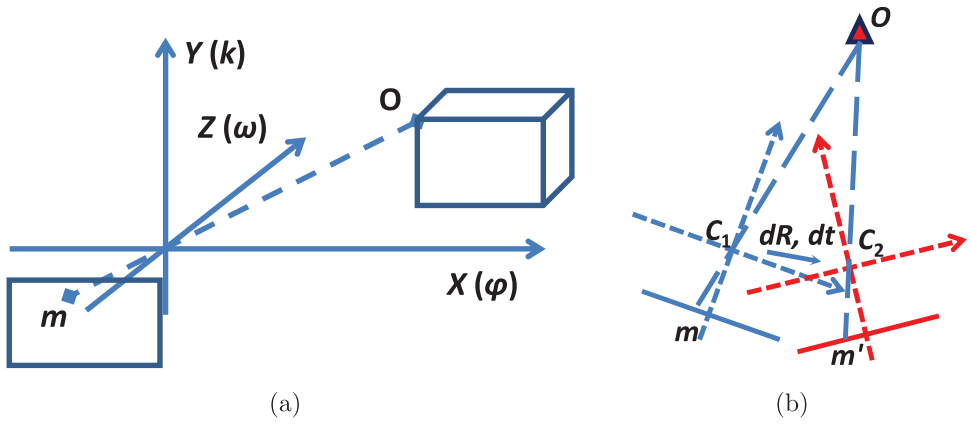

Fig. 4. (a) Indicates the coordinate system assuming the viewpoint as the origin, and (b) indicates rotation and translation errors between the real-world viewpoint and the measured viewpoint.

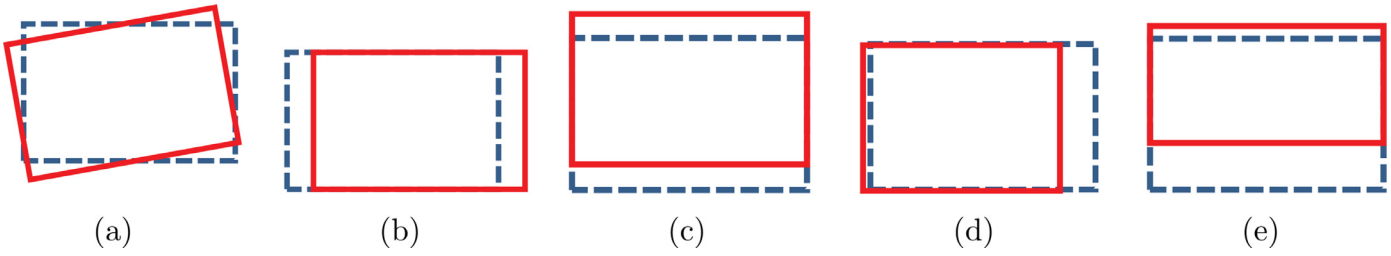

Fig. 5. The relationship between input error correction and output registration compensation. (a) Rotation correction along Z axis, (b) rotation correction along Y axis, (c) rotation correction along X axis, (d) horizontal correction along X axis, and (e) horizontal correction along Y axis.

4.3. Judging conditions and strategies for interactive process

In fact, users can hardly distinguish if the error is induced by the positioning error or orientation error as shown in Fig. 2, which raises difficulties in practical operations. Therefore, it is necessary to discuss how to distinguish these two types of errors in order to lay a theoretical foundation and guidance for precise registration.

Since the wireframe models cannot perfectly match the real-world objects, instead, either a vertical or horizontal bottom edge of the wireframe model can be selected for matching. Therefore, the process of precise registration can be implemented by panning the viewpoint so as to let such vertical or horizontal bottom edge match the corresponding edge of real-world objects.

If a linear function of the edge on the projection 2D surface is expressed as $y = kx + b$, $k = dy/dx$, in accordance to Eqs. (2)–(7), $k = dm'_y/dm'_x$ and the corresponding values of k can be calculated as follows:

$$k_2 = -\frac{X}{Y}, \quad k_3 = 0, \quad k_4 = \infty, \quad k_5 = 0, \quad k_6 = \infty, \quad k_7 = \frac{Y}{X} \tag{9}$$

where k_2–k_7, respectively, correspond to Eqs. (2)–(7). For example, suppose $X = Y = 10$ m, then $k_2 = -1$ and $k_7 = 1$.

It can be inferred from these k values that: when inputting e_ω in Eq. (2), horizontal or vertical edge can be improperly rotated since the vertices of edges at different locations are affected by different compensations; when inputting e_k in Eq. (3) and e_φ in Eq. (4), vertical and horizontal edges can be adjusted by horizontal and vertical compensations, respectively. When inputting X and Y positioning corrections, the obtained compensations are related to the distance between the target object and the observer. Thus the horizontal or vertical edge can be adjusted by corresponding compensation, where the value of compensation is inversely proportional to the distance between the edge and the user. Rotation compensation can be raised when inputting the positioning correction along Z axis. However, comparing with other compensations, such rotation compensation can be ignored. These relationships are further explained in Fig. 5. Only these five relationships play a role in the basic relationships, other relationships can be considered as the combination of some of these basic relationships.

Based on the above discussion, the conditions that require manual interactive process are concluded as follows:

Condition 1: If there exists an offset illustrated in Fig. 5(a), it is mainly caused by the orientation error along Z axis.
Condition 2: If there exists an offset between two vertical edges, it is mainly caused by the error of panning along Y axis; if there is offset between two horizontal edges, it is mainly caused by the error of panning along X axis.
Condition 3: If there exists homogeneous offset formed in vertical or horizontal direction, it is mainly caused by the angle-measurement error along X or Y axis.

In practice, it is hard to distinguish the offset between Conditions 1 and 2. The best way to tackle this problem is to use Real Time Kinematic Global Positioning System (RTKGPS) to reduce the measurement error up to centimeter level, and then the offset caused by positioning error can be ignored. In this case, since the orientation error along the three axes have different offset patterns, the corresponding corrections can be inputted to improve the registration accuracy.

When a lowly accurate (consumer-grade) GPS is used, a highly accurate IMU can be acquired to minimize the orientation error along X and Y axes, therefore only the error along Y axis (i.e., the error of north) needs to be considered. Assuming there is no significant change in elevation, the positioning error along Y axis can be ignored. In this case, the offset illustrated in Fig. 5(b) and (d) should be considered. When these two measurement errors are

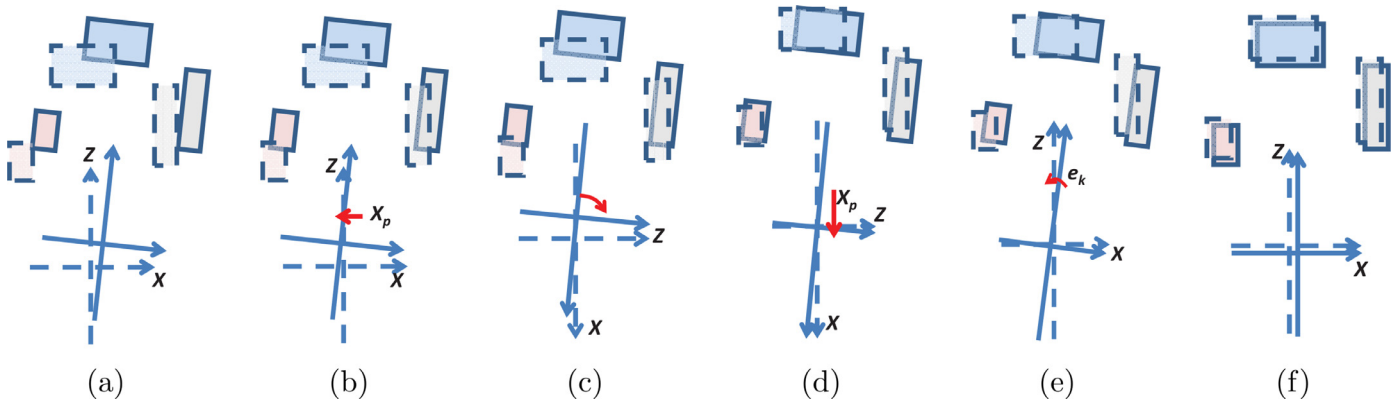

Fig. 6. An illustration of steps for correcting both rotation error and translation error (dot lines express the real-world position of the target and the viewpoint, solid lines express the location with error). (a) Before correction, (b) inputting horizontal offset X_p, (c) rotation view direction by 90°, (d) inputting horizontal offset X_p once more, (e) rotation view direction back by 90°, and inputting rotate offset e_k, and (f) the final registration result.

found, the total offsets can be approximately described as

$$dm'_x = f(e_k + \frac{X_p}{Z}) \tag{10}$$

where e_k refers to the orientation parameter along Y axis and X_p refers to the positioning parameter along X axis. Eq. (10) indicates that positioning offset depends on distance, while angle offset has nearly no relationship with distance. Consequently, certain correction strategies are proposed as follows:

Strategy 1: When correction starts, user stays at any location with a straight observing direction ahead. A filtering algorithm can be used to stabilize the positioning and orientation measurement data.

Strategy 2: The sequence of inputting rotation corrections should start with the visible objects located at the farthest to the nearest.

Strategy 3: If there is no object which is far enough, i.e., nearer than a specific distance Z, the vertical or horizontal offset is greater than the rotation offset. In this case, the vertical or horizontal correction should be inputted first. Eq. (5) can be used to compute Z. For example, if the tolerable registration error is 1 pixel, i.e., $dm'_x \le 1$, then $fX_p/Z = dm_x \le 1 \Rightarrow fX_p \le Z$. Suppose $f = 1000\,p$, the accuracy of GPS is 5 cm, then $Z \ge 1000 \times 0.05\text{ m} = 50\text{ m}$.

Strategy 4: Select two vertical views and input horizontal offset correction twice, the horizontal positioning measurement error along X and Y axes are subsequently corrected.

Strategy 5: By inputting the vertical correction, the vertical offsets of two views are adjusted to be identical, and then input the rotation correction.

Strategy 6: The positioning and angular corrections cannot go beyond the maximum error values of sensors.

4.4. Algorithm for input compensation of gyroscope

The error of gyroscope changes over time and the rotation error along Y axis usually increases progressively over the system running time, thus the compensation required to be inputted during the interactive registration process raises dramatically. On one hand, the correction process would be delayed. On the other hand, it is hard to distinguish the positioning error during the registration process for those objects close to the observer. In order to make input rotation correction to be smaller to τ, suppose the ith inputted correction is k_i, then

$$k_i = k_{i-1} + \sigma(t)\tau \tag{11}$$

where τ comes from priori knowledge (e.g., the value of 0.5°), $\sigma(t)$ refers to an function related to time that can be described as

$$\sigma(t) \approx Int\left\lfloor \frac{k_i - k_{i-1}}{\tau} \right\rfloor \tag{12}$$

By using Eqs. (11) and (12), the inputted corrections can be accumulated, which enables systematic error of gyroscope to be stably and easily corrected by users. For example, users can first correct the error of rotating from north of gyroscope with the reference of 3D GIS before conducting registration.

4.5. Steps of interactive registration

The specific correction process with consideration of both positioning and orientation error is illustrated in Fig. 6. When the accuracy of RTKGPS reaches to centimeter level, only steps e and f are required to accomplish the correction. More specifically, the interactive registration can be conducted based on the following steps:

1. Before turning on the mobile terminal, determine the accuracy range of GPS and gyroscope being used. The values of the accuracy range are inputted into the system with reference to Strategy 6.
2. GPS and gyroscope are utilized to receive the positioning and orientation data. Coarse registration is sequentially conducted between the virtual and physical environments.
3. Based on Strategy 1, precise registration starts to control the measurements from positioning and orientation sensors.
4. Following Conditions 1 and 3, the user performs corrections sequentially based on Strategies 2 and 5. In complex environment, Strategy 2–5 might be repeated in two to three times (refer Fig. 6(b)–(f)).
5. The coefficient of input correction of gyroscope is adjusted and updated based on Eq. (12).

5. Experiment and discussion

Based on the system architecture and registration method proposed, a prototype of ARGIS was developed and several field tests were preformed. Some issues about the proposed system are discussed in the following subsections.

5.1. Understanding the linear relationship between the input correction and output compensation

Since the relationship between Eqs. (3) and (4) is linear only under certain situations, such situations need to be figured out for

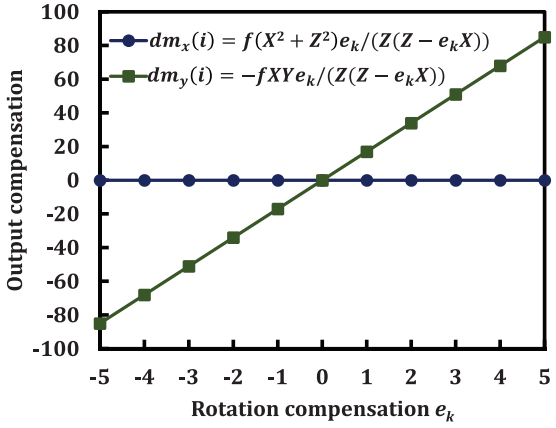

Fig. 7. The relationship between the rotation compensation e_k of Y and the projected compensations dm_x and dm_y.

determining the adopted sensor accuracy. Only Eq. (3) is selected to be further discussed since both equations are similar.

Suppose an object is placed in front of a viewpoint, and the position from the viewpoint is located at $Z = 100$ m, $X = 10$ m, $Y = 10$ m. If the focal length, f, is set to be 960 pixels and the input rotation compensation, e_k, is bounded by -5° and 5°, the output compensation changes in a linear manner (Fig. 7). In our experiment, MTi sensors whose dynamic orientation accuracy is reported as 2° was used, therefore it can guarantee that a linear relationship exists between the rotation compensation e_k of Y and the projected compensations dm_x and dm_y during the interactive registration process.

5.2. The ARGIS prototype

The prototype was tested on the campus of Peking University, Beijing, China. A 1:5000 campus map was used to generate the 3D building models stored in the 3D GIS database with an assumption of identical building height. In order to distinguish different buildings, the 3D models were rendered in different colors (Fig. 8). The buildings' rooftop is not taken into account to perform interactive registration. Instead, only vertical edges and horizontal footprints are used. The elevation data of GPS was ignored in this experiment due to the unavailability of DEM data. However, the elevation correction can be still inputted to match the bottom lines of virtual and real-world buildings during the interactive registration process. Since the compensation of elevation and orientation are independent with each other, the interactive operation can be conducted efficiently.

Fig. 9 shows the ARGIS prototype consisting of the two mobile terminals. The orientation sensors are the MTi products manufactured by XSens (https://www.xsens.com/products). The static accuracy of yaw is less than 1.0° and the accuracy of both roll and pitch is less than 0.5°. A RTKGPS was equipped in the mobile terminal as shown in Fig. 9 (left). Through receiving differential signals of base stations, the horizontal positioning accuracy obtained by the RTKGPS could yield an accuracy of centimeter level. Since the high accurate GPS sensor was equipped, only the measurement error of MTi needed to be calibrated during the interactive registration. The handheld devices for the interactive operations, including the devices comprise of a type of multi-channel game controller and a mini keyboard, are illustrated in Fig. 9. Input correction of the position and orientation could be achieved through using different channels of the game controller.

Moreover, another mobile terminal, as shown in Fig. 9 (right), was equipped with another GPS sensor (MTiG) having 2-m positioning accuracy as reported. Although this equipment was much lighter and more portable, the efficiency of interactive registration was much lower than the other system equipped with RTKGPS during the practice. In case of small view field, an ideal result of registration can come out after conducting several rounds of trials. Actually, the MTi-G sensor was mainly used for evaluating the result conducted by RTKGPS. In practice, RTK module is strongly recommended to be equipped since it is becoming smaller and lighter and cheaper.

Fig. 10 refers to the scenes shown in the head mounted display (HMD) before and after conducting precise registration. More specifically, the colored objects indicate the 3D wireframe models of buildings with annotation from Fig. 8. Fig. 10(a) refers to the scene shown in HMD before the correction. It is apparent that there exists a gap (between the two yellow lines) between the virtual object (the pink box) and the corresponding target building in real world. Fig. 10(b) indicates the scene shown in the HMD after the correction. The edge of the building is aligned to the edge of the virtual object (refer the edge in the yellow box). Fig. 10(c) and (d) shows the results of interactive registration of the other two buildings.

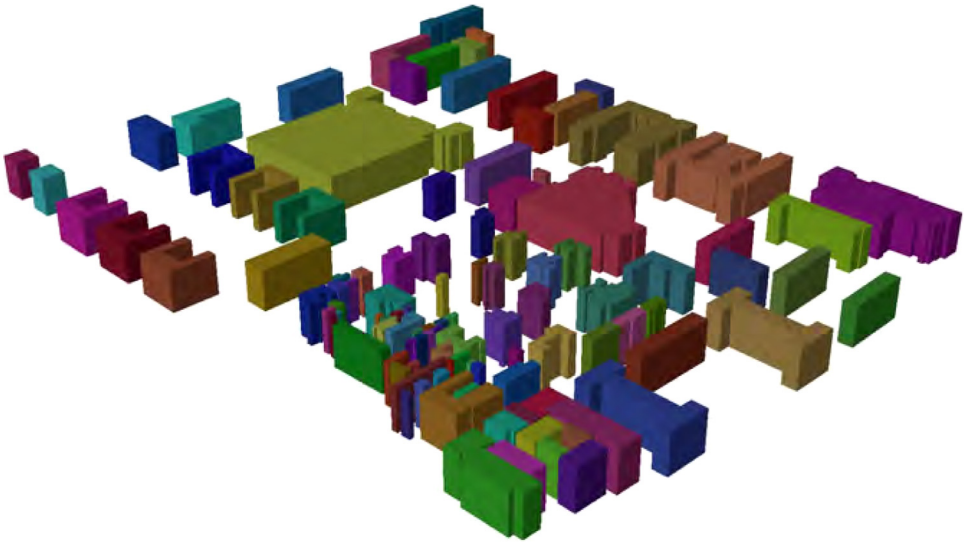

Fig. 8. The building wireframe models of Peking University campus.

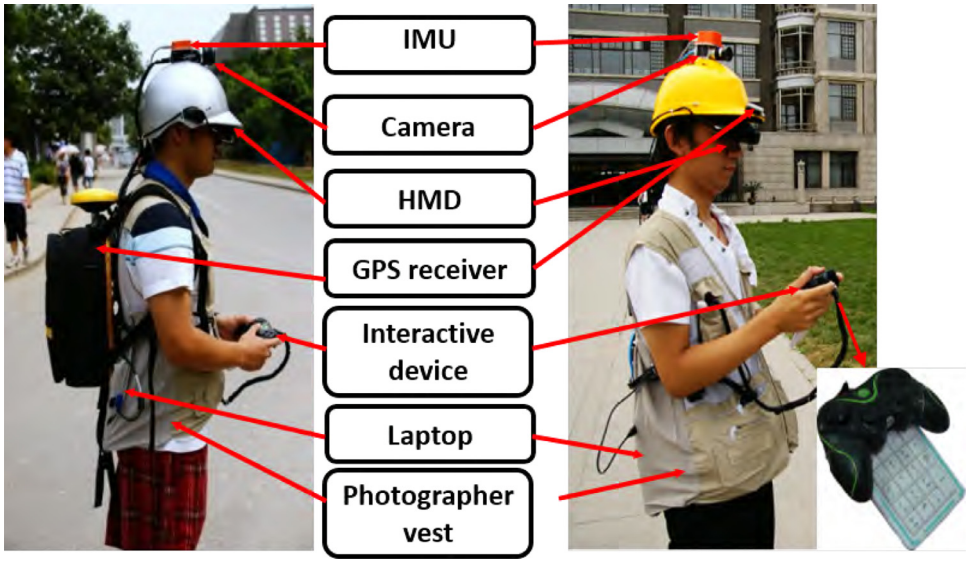

Fig. 9. Mobile terminals of ARGIS prototype. The mobile terminal shown in the left hand side uses RTKGPS to receive positioning data, where the horizontal accuracy is less than 5 cm. The mobile terminal shown in the right hand side uses MTi-G to receive positioning data, where the horizontal accuracy is within 5 m. The right-bottom corner shows the interactive device.

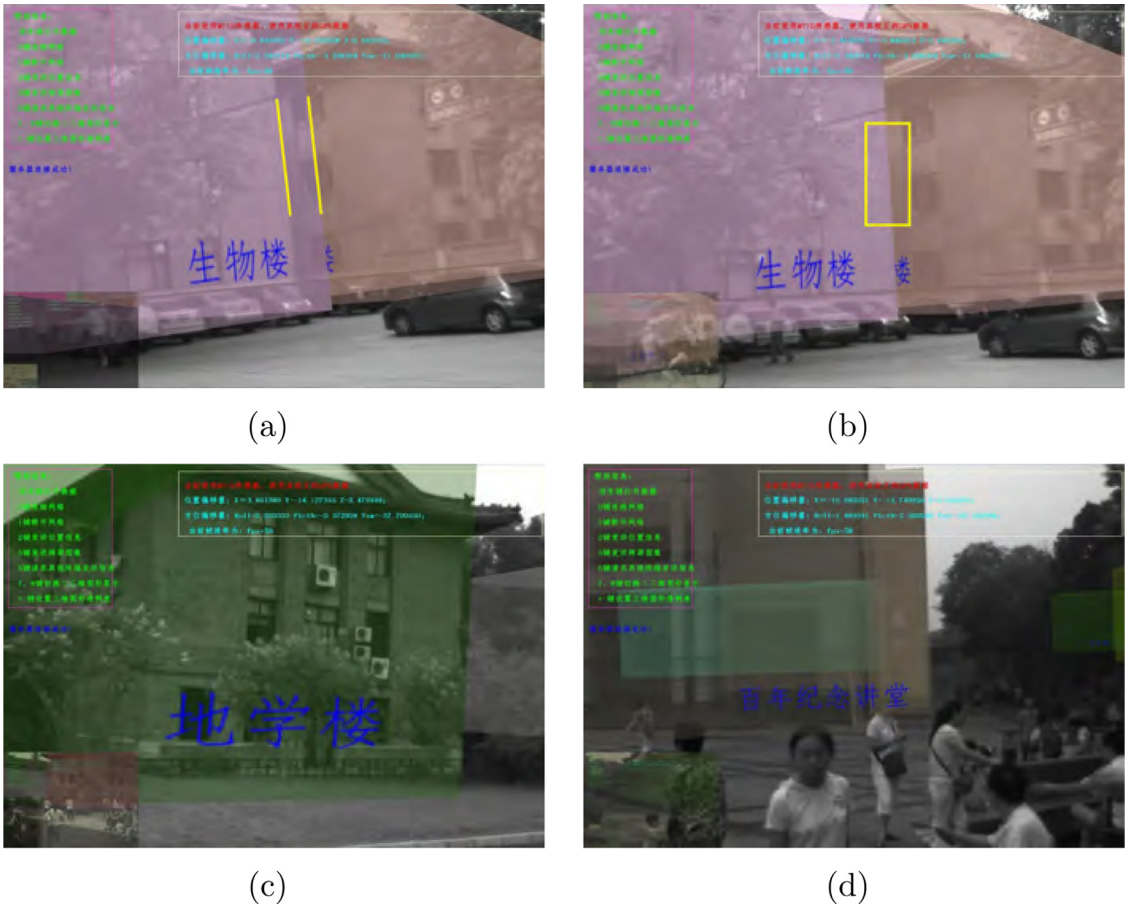

Fig. 10. Selected scenes shown in the HMD during the interactive registration (the 3D building models and annotations are registered to the buildings on the campus) (For interpretation of the references to color in this figure text, the reader is referred to the web version of this article.).

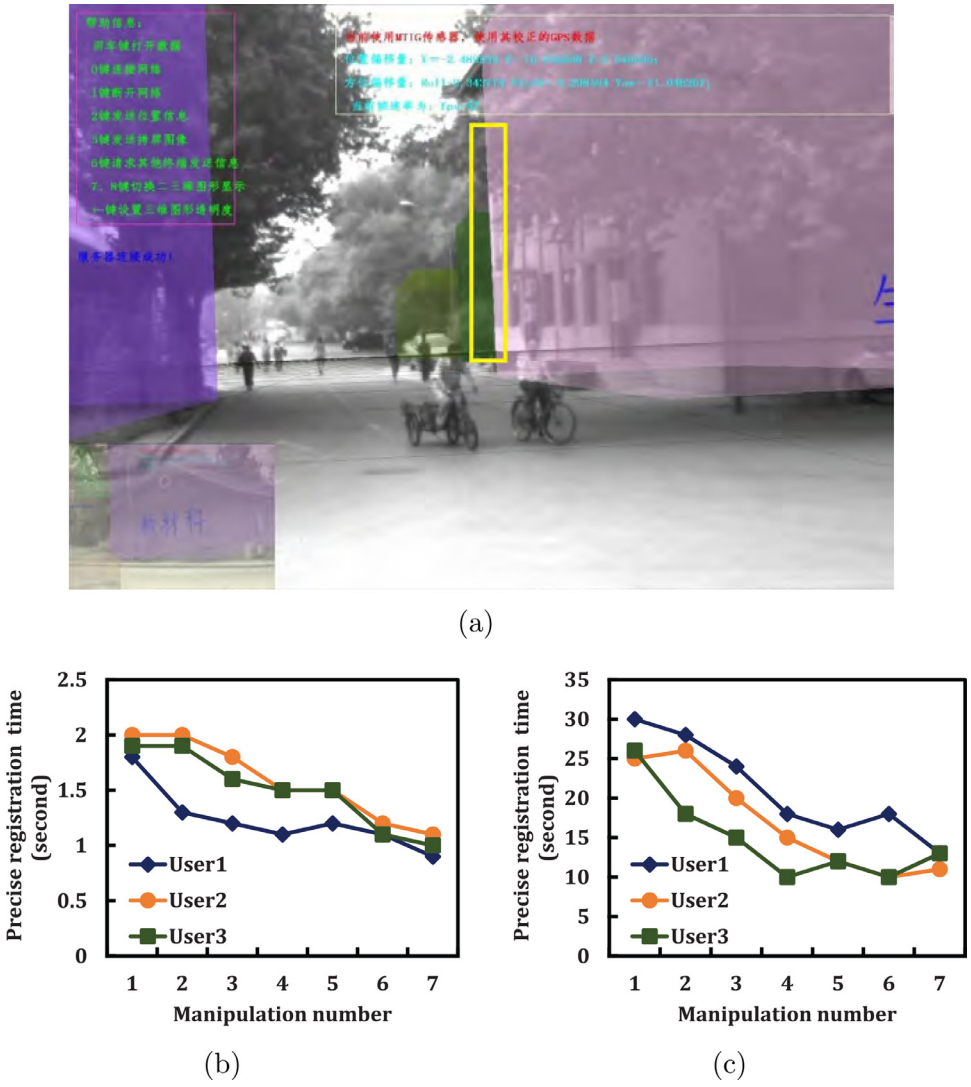

Fig. 11. The efficiency of precise registration. The target registration result (a); Three users conducted precise registration seven times using RTKGPS (b); Three users conducted precise registration using low accuracy GPS receivers (MTi-G) seven times (c). (For interpretation of the references to color in this figure text, the reader is referred to the web version of this article.)

5.3. Discussion

5.3.1. Data accuracy in 3D GIS

As aforementioned, only the 1:5000 campus map was used to create the 3D wireframe models with an assumption of identical building height in 3D GIS. This type of data works well for matching the building outlines in the physical environment on a flat ground surface. If the system works in rugged terrain, high accurate DEM are needed to be stored in the 3D GIS database.

In this study, buildings were mainly selected to be the data reference source for precise registration. Thus, the system is ideal to work in urban area. Yet, some specific target objects need to be identified in the corresponding 3D GIS in advance if recognizable target objects are lacking. With the rapid growth of satellite remote sensing and laser scanning technology (Shaker, Yan, & Easa, 2010, 2011; Yan, Shaker, & El-Ashmawy, 2015), 3D city models can be easily extracted with high precision. In addition, there have been some attempts to extract 3D geo-positioning information from panoramic camera data at street level (Tsai & Chang, 2013; Yan, Shaker, & Easa, 2013). With the aid of these techniques to generate 3D wireframe models, the registration error of ARGIS can be significantly reduced with the improved 3D GIS data quality.

5.3.2. Efficiency of interactive registration

As discussed before, in case of using RTKGPS, interactive registration can be efficiently manipulated. When using lowly accurate GPS receiver (MTi-G), the registration offsets of positioning and orientation measurements have to be corrected multiple times. Users can easily accomplish the interactive registration if they have prior knowledge about the registration mechanism and become more familiar with the interactive devices. In our experiment, a specific building was selected as a reference to sequentially manipulate the interactive registration seven times by three different users using RTKGPS and lowly accurate GPS receiver (MTi-G), respectively. Fig. 11(a) illustrates the targeted result for each manipulation of registration. The users were only demonstrated how to use the interactive device to input corrections to adjust the virtual objects. The result reports that the time each user spent on to align the vertical edge of the building, shown in the yellow box in Fig. 11(a), decrease while number of manipulation increases (Fig. 11(b) &(c)). In general, all the three users were able to finish the precise registration within 1–2 s using RTKGPS, while 10–30 s using MTi-G GPS receivers. Apparently, it is much more efficient to accomplish the registration by using RTKGPS than GPS with low accuracy.

5.3.3. Difference between ARGIS and traditional AR System

Most of existing AR systems focus on the precise matching between virtual objects and real-world objects by using vision tracking method. In other words, they either cannot work in the environment with objects which are hard to be tracked or need to predefine specific markers to be tracked. Therefore, it is hard to implement them in outdoor environment without any predefined markers, and it is not realistic to place artificial markers in outdoor environment every time before starting outdoor AR system. Although the vision-based AR system has a very good performance on registration in indoor environment or marked outdoor environment, if all other applications and researches strictly pursue this target, it will raise certain limits for the development of AR system, particularly in outdoor environment.

In contrast, in some fields, such as GIS field, users do not care about how accurate virtual objects and real-world objects are matched with each other; instead, users would prefer to pay more attention on the information of which they are interested in, such as the attributes and structure of the objects. To achieve this goal, the wireframe models and annotated information are used as augmented information in this study, any other attribute information in 3D GIS can also be incorporated. More importantly, in this study, the purpose of precise registration aims to obtain the precise geolocation of the virtual objects in the physical outdoor environment without using vision tracking, which helps to further explore other outdoor AR applications (e.g., utility underground survey and emergency rescue).

The other significant difference between ARGIS and traditional AR system is that ARGIS supports on-line users to communicate with each other, even the server side is allowed to communicate with the front-end, while traditional AR system is designed for off-line applications. In this case, the proposed interactive device for the precise registration plays a key role in inputting certain simple commands for a specific communication. For example, in the experiment, each user can input “5” to request augmented views of the scene from other on-line users and browse them in a sub-window located at the bottom left corner of the main interface.

6. Conclusions

Registering virtual objects to the corresponding real objects is essential for AR system. However, most of the existing methods for registration are based on vision tracking, which is hard to be implemented for outdoor AR system. To tackle this issue, this paper presents a new registration mechanism for outdoor AR system by combining 3D GIS and AR system, named as ARGIS. Unlike traditional AR system, the proposed ARGIS does not require the use of vision tracking method for registration, and instead, a handheld device is proposed to be used to perform precise registration with referencing the 3D wireframe models built in the 3D GIS database. The experimental results show that users are able to finish the precise registration within 2 s by using RTKGPS and 30 s by using lowly accurate GPS receivers, which is significantly faster than that of using vision tracking method. Moreover, the proposed interactive device for ARGIS also provides a convenient tool for the communication among different on-line users.

The outcome of this study can contribute to bring GIS applications to outdoor environment and increase the feasibility of AR system working in outdoor environment. However, the main limitation of the proposed system is the power supply that shares among camera, HMD, laptop, and thus raise computational burden resulting in an intensive power consumption. In the future, the system is planned to be tested in high performance smart devices, which is expected to develop a low-cost and more portable platform.

Acknowledgments

This work has been funded by the National Key Technology R&D Program of China (grant no. 2012BAH27B02). The authors would like to thank Dr. Wai-Yeung Yan for his help in developing the manuscript, and the two anonymous reviewers for their constructive comments, which helped improve the paper.

References

Azuma, R., Baillot, Y., Behringer, R., Feiner, S., Julier, S., & MacIntyre, B. (2001). Recent advances in augmented reality. *IEEE Computer Graphics and Applications, 21*(6), 34–47.

Azuma, R., Hoff, B., Neely III, H., & Sarfaty, R. (1999). A motion-stabilized outdoor augmented reality system. In *Proceedings of the 1999 IEEE virtual reality* (pp. 252–259). IEEE.

Azuma, R., Neely, H., Daily, M., & Leonard, J. (2006). Performance analysis of an outdoor augmented reality tracking system that relies upon a few mobile beacons. In *Proceedings of the fifth IEEE and ACM international symposium on mixed and augmented reality* (pp. 101–104). IEEE Computer Society.

Azuma, R. T., et al. (1997). A survey of augmented reality. *Presence, 6*(4), 355–385.

Behringer, R., Park, J., & Sundareswaran, V. (2002). Model-based visual tracking for outdoor augmented reality applications. In *Proceedings of the 2002 international symposium on mixed and augmented reality (ISMAR)* (pp. 277–322). IEEE.

Behzadan, A. H., & Kamat, V. R. (2007). Georeferenced registration of construction graphics in mobile outdoor augmented reality. *Journal of Computing in Civil Engineering, 21*(4), 247–258.

Carmigniani, J., Furht, B., Anisetti, M., Ceravolo, P., Damiani, E., & Ivkovic, M. (2011). Augmented reality technologies, systems and applications. *Multimedia Tools and Applications, 51*(1), 341–377.

Feiner, S., MacIntyre, B., Höllerer, T., & Webster, A. (1997). A touring machine: Prototyping 3D mobile augmented reality systems for exploring the urban environment. *Personal Technologies, 1*(4), 208–217.

Fong, W., Ong, S.-K., & Nee, A. Y. (2009). Computer vision centric hybrid tracking for augmented reality in outdoor urban environments. In *Proceedings of the eighth international conference on virtual reality continuum and its applications in industry* (pp. 185–190). ACM.

Ghadirian, P., & Bishop, I. D. (2002). Composition of augmented reality and GIS to visualize environmental changes. In *Proceedings of the joint Aurisa and institution of surveyors conference* (pp. 25–30). CiteSeer.

Jiang, B., Neumann, U., & You, S. (2004). A robust hybrid tracking system for outdoor augmented reality. In *Proceedings of the IEEE virtual reality conference*. Chicago, Illinois, USA: March, IEEE.

King, G. R., Piekarski, W., & Thomas, B. H. (2005). Arvino-outdoor augmented reality visualisation of viticulture GIS data. In *Proceedings of the fourth IEEE and ACM international symposium on mixed and augmented reality.* (pp. 52–55). IEEE.

Lee, G. A., Dunser, A., Nassani, A., & Billinghurst, M. (2013). Antarcticar: An outdoor AR experience of a virtual tour to antarctica. In *Proceedings of the 2013 IEEE international symposium on mixed and augmented reality-arts, media, and humanities (ISMAR-AMH),* (pp. 29–38). IEEE.

Lee, G. A., Dunser, A., Kim, S., & Billinghurst, M. (2012a). CityViewAR: A mobile outdoor ar application for city visualization. In *Proceedings of the 2012 IEEE international symposium on mixed and augmented reality (ISMAR-AMH)* (pp. 57–64). IEEE.

Lee, J.-D., Huang, C.-H., Huang, T.-C., Hsieh, H.-Y., & Lee, S.-T. (2012b). Medical augment reality using a markerless registration framework. *Expert Systems with Applications, 39*(5), 5286–5294.

Li, Y., Weng, D., Zhou, H., Hao, J., & Zhao, L. (2013). Kaidan: An outdoor AR puzzle adventure game. In *Proceedings of the 2013 IEEE international symposium on mixed and augmented reality-arts, media, and humanities (ISMAR-AMH)* (pp. 7–11). IEEE.

Liarokapis, F., Greatbatch, I., Mountain, D., Gunesh, A., Brujic-Okretic, V., & Raper, J. (2005). Mobile augmented reality techniques for geoVisualisation. In *Proceedings of the 2005 ninth international conference on Information visualisation* (pp. 745–751). IEEE.

Mok, E. C., Shea, G. Y., & Yan, W. Y. (2004). Geolocation positioning with wireless cellular network in Hong Kong. *Hong Kong Surveyor, 15*(2), 23–30.

Piekarski, W., & Smith, R. (2006). Robust gloves for 3D interaction in mobile outdoor AR environments. In *Proceedingsof the IEEE/ACM international symposium on mixed and augmented reality (ISMAR)* (pp. 251–252). IEEE.

Piekarski, W., & Thomas, B. H. (2002). *Unifyingaugmented reality and virtual reality user interfaces*. Adelaide: University of South Australia.

Pun-Cheng, L. S., Mok, E. C., Shea, G. Y., & Yan, W. Y. (2007). EASYGO – A public transport query and guiding LBS. In *Location based services and telecartography* (pp. 545–556). Springer.

Reitmayr, G., Eade, E., & Drummond, T. (2005). Localisation and interaction for augmented maps. In *Proceedings of the fourth IEEE/ACM international symposium on mixed and augmented reality* (pp. 120–129). IEEE Computer Society.

Satoh, K., Anabuki, M., Yamamoto, H., & Tamura, H. (2001). A hybrid registration method for outdoor augmented reality. In *Proceedings of the IEEE and ACM international symposium on augmented reality.* (pp. 67–76). IEEE.

Schall, G., Mendez, E., Kruijff, E., Veas, E., Junghanns, S., Reitinger, B., & Schmalstieg, D. (2009a). Handheld augmented reality for underground infrastructure visualization. *Personal and Ubiquitous Computing, 13*(4), 281–291.

Schall, G., Wagner, D., Reitmayr, G., Taichmann, E., Wieser, M., Schmalstieg, D., & Hofmann-Wellenhof, B. (2009b). Global pose estimation using multi-sensor fusion for outdoor augmented reality. In *Proceedings of the eighth IEEE international symposium on mixed and augmented reality (ISMAR)* (pp. 153–162). IEEE.

Seo, D. W., & Lee, J. Y. (2013). Direct hand touchable interactions in augmented reality environments for natural and intuitive user experiences. *Expert Systems with Applications, 40*(9), 3784–3793.

Shaker, A., Yan, W. Y., & Easa, S. (2010). Using stereo satellite imagery for topographic and transportation applications: An accuracy assessment. *GIScience & Remote Sensing, 47*(3), 321–337.

Shaker, A., Yan, W. Y., & Easa, S. (2011). Construction of digital 3D highway model using stereo IKONOS satellite imagery. *Geocarto International, 26*(1), 49–67.

Thomas, B., Close, B., Donoghue, J., Squires, J., De Bondi, P., Morris, M., & Piekarski, W. (2000). ARQuake: An outdoor/indoor augmented reality first person application. In *Proceedings of the fourth international symposium on wearable computers* (pp. 139–146). IEEE.

Tsai, V. J., & Chang, C.-T. (2013). Three-dimensional positioning from Google street view panoramas. *IET Image Processing, 7*(3), 229–239.

Uchiyama, H., Saito, H., Servières, M., & Moreau, G. (2009). AR city representation system based on map recognition using topological information. In *Virtual and mixed reality* (pp. 128–135). Springer.

Vincent, T., Nigay, L., & Kurata, T. (2013). Handheld augmented reality: Effect of registration jitter on cursor-based pointing techniques. In *Proceedingsof the 25 IÈME conférence francophone on l'interaction homme–machine* (pp. 1–6). ACM.

Yan, W. Y., Shaker, A., & Easa, S. (2013). Potential accuracy of traffic signs' positions extracted from Google street view. *IEEE Transactions on Intelligent Transportation Systems, 14*(2), 1011–1016.

Yan, W. Y., Shaker, A., & El-Ashmawy, N. (2015). Urban land cover classification using airborne LiDAR data: A review. *Remote Sensing of Environment, 158*, 295–310.

Yovcheva, Z., Buhalis, D., & Gatzidis, C. (2012). Smartphone augmented reality applications for tourism. *e-Review of Tourism Research (eRTR), 10*(2), 63–66.

Yuan, M., Ong, S.-K., & Nee, A. Y. (2006). Registration using natural features for augmented reality systems. *IEEE Transactions on Visualization and Computer Graphics, 12*(4), 569–580.

Zlatanova, S., & van den Heuvel, F. (2002). 3D GIS for outdoor ar applications. In *Proceedings of the third international symposium on mobile multimedia systems & applications, December, Delft, the Netherlands* (pp. 117–124).