Outdoor See-Through Vision Utilizing Surveillance Cameras

Yoshinari Kameda[1], Taisuke Takemasa, and Yuichi Ohta[2]
Department of Intelligent Interaction Technologies,
Graduate School of Systems and Information Engineering,
University of Tsukuba
1-1-1 Tenoudai, Tsukuba, 305-8573, Japan
kameda@ieee.org[1], ohta@acm.org[2]

Abstract

This paper presents a new outdoor mixed-reality system designed for people who carry a camera-attached small handy device in an outdoor scene where a number of surveillance cameras are embedded. We propose a new functionality in outdoor mixed reality that the handy device can display live status of invisible areas hidden by some structures such as buildings, walls, etc. The function is implemented on a camera-attached, small handy subnotebook PC (HPC). The videos of the invisible areas are taken by surveillance cameras and they are precisely overlapped on the video of HPC camera, hence a user can notice objects in the invisible areas and see directly what the objects do. We utilize surveillance cameras for two purposes. (1) They obtain videos of invisible areas. The videos are trimmed and warped so as to impose them into the video of the HPC camera. (2) They are also used for updating textures of calibration markers in order to handle possible texture changes in real outdoor world. We have implemented a preliminary system with four surveillance cameras and proved that our system can visualize invisible areas in real time.

1. Introduction

Outdoor mixed reality is a new technology to enrich information environments for people in outdoor scenes. There are various applications including pedestrian navigation[28][19] that expect the development of outdoor mixed-reality technology.

In this paper, we propose a new mixed-reality system with which users can see through buildings on streets so that they can check hidden areas based on mixed-reality technology. As the device should be portable and small enough to carry, we implemented the see-through function on a camera-attached, small handy subnotebook PC (HPC).

One of the advantages of our approach is that users can see the areas occluded by buildings or walls in real time by utilizing cameras embedded in the environment for surveillance purposes. In order to let users recognize objects in the occluded areas and understand how they move, visual information of the occluded areas should be precisely superimposed over the video of HPC camera at HPC display. Visual information to be integrated with a video image of the HPC camera consists of two data sources; one is CG objects based on CAD data models that represents static objects in the areas, and the other is live videos taken at remote viewpoints that visualize objects in the hidden areas.

Displaying just a raw video of a remote camera on HPC display is not an effective approach so as to let users instantly understand the situation of the hidden areas because it takes time to recognize the location and orientation of the remote camera towards the hidden areas just by watching the video image even with the support of a 2D map on which the remote camera is marked.

In outdoor scenes, a HMD is not the only solution that offers MR environment. We adopt a camera-attached HPC , which is similar to PDA or handheld PC. We think there are three reasons to use a HPC for outdoor mixed reality. One is that it is feasible to mount all the sensors including a camera, a GPS, and an inertia sensor together with a display onto one small solid package. The second reason is that it is not recommended to cover field of view of users by any device that cannot be detached easily because of security reasons in outdoor scenes. The third reason is that it is not important to align the user's viewpoint against HPC display in which the video of HPC camera and CG objects are mixed. In outdoor scenes, the objects to be mixed are usually at least several meters away

Proceedings of the Third IEEE and ACM International Symposium on Mixed and Augmented Reality (ISMAR 2004)

from a user. At this range, the difference between the viewpoint of the user and that of the HPC camera is not critical to recognize the mixed world on the HPC display.

In this approach, we assume that the CAD data of some buildings or structures in a scene can be accessible and a number of fixed surveillance cameras are embedded in a scene. The surveillance cameras should be calibrated against the world in advance.

Another advantage of our approach is the use of pre-existing structures as calibration markers. Calibration markers are usually effective to estimate the location and orientation of a camera, but it is not feasible to place special markers in a wide outdoor scene. Therefore, we focus on distinct substructures of buildings and use them as markers. We call them landmarks in this paper.

When a video image of a surveillance camera is imposed on a HPC camera image, the homography matrix between the two cameras should be obtained. A straight-forward approach is to estimate the matrix directly by referring image features that can be observed by both cameras. However, in outdoor scenes, overlap of the view volumes of two cameras is usually small so that the estimation of the matrix will be unstable and inaccurate.

Another approach is to decompose the estimation process into two phases. The first phase is to estimate camera projection matrix against the world coordinate system and the second phase is to find a match between the world coordinate system and CAD model coordinate system. In this approach, the video segment of a surveillance camera is projected on a HPC image by using the fixed relation between the surveillance camera and the world. Various on-line camera registration methods have been proposed [7][12][26][32][2] for the first phase. If special calibration markers that are registered in the CAD model can be installed in the scene, camera registration would be achieved by these methods. However, it is not feasible to set artificial markers in a wide outdoor scene. Hence, we use substructures of buildings instead of artificial markers.

The other advantage of our approach is utilization of surveillance cameras for obtaining the latest textures of landmarks.

When we use landmarks as calibration markers, what make image processing difficult are environmental changes across time, e.g. possible variations of color and/or shape of the landmarks. These variations may be caused by light condition changes, color fading, aging, human factors, etc.

Therefore, we propose to observe the landmarks by surveillance cameras and obtain the latest textures of the landmarks when they are requested. The appearance difference between the landmarks and their images taken by HPC camera can be diminished because the appearances of the landmarks in the CAD model are always the latest ones.

The rest of the paper is organized as follows. In section 2, recent advances of augmented reality technologies that could be applied to outdoor mixed reality are discussed. Then, our proposed method is described in section 3. Camera registration procedure, that ensures accurate mixed reality on HPC device, is shown in section 4 and the visualization of invisible area is described in section 5. We show experimental results in section 6 and conclude the paper in section 7.

2. Related Works

In some recent and advanced applications, a PDA or a handheld PC, which uses small displaying equipment similar to our HPC, is used to display annotations that are subject to the location of a user [16]. While annotation information is shown in 2D maps, texts, and/or images on these devices, it is not required to estimate accurate location and orientation hence GPS and beacons[21] are sufficient for these applications. Except for visualizing the live status of hidden areas, Pasman and Woodward presented a system design of PDA-based MR application [18] and their PDA can display 3D CG objects aligned with a PDA camera.

Advanced Global Positioning Systems (GPS) including differential GPS and real-time kinematic GPS are considered to be promising to estimate positions in outdoor scenes in general. As for pedestrians, sometimes urban canyon problem is very serious because they may come closer to buildings than vehicles. To enhance the usability of GPS, map information and trajectory of user are used together by Cui and Ge[8].

One of the solutions to compensate the difficulties of GPS is to utilize other sensors including magnetometer, accelerometers, etc.

Kourogi and Kurata proposed a walker navigation system[13][14]. Tenmoku et al. proposed a wearable augmented reality system that shows annotative tags of objects seen through HMD in outdoor scenes[28]. In these applications, precise alignment of tags against real world is not requested since the tags indicate direction of objects to visit or annotations of objects in the scene. Therefore, they could realize the tag display by using only non-vision sensors that cannot perform accurate estimation for geometric overlay.

Despite recent advances of positioning with GPS, magnetometers, and accelerometers, they could not achieve accurate measurement of positions and orientations to realize mixed reality on a HPC display in outdoor scenes, because overlapping HPC camera images with CG models and warped videos taken from surveillance cameras needs more accurate estimation. Therefore, image-based camera registration is necessary.

One of the major approaches for camera registration is to place artificial markers. In outdoor scenes, it is not practical to place explicit markers though they are useful to perform accurate camera registration in indoor environment [31] [32] [2].

Kumar et al. proposed a 3D manipulation method that can be used both in markerless camera registration and scene recovery[17]. Chia et al. also proposed a camera registration method that only uses natural features in a scene[6]. They work well when a sufficient number of feature points are found in video sequence. Comport et al. proposed a robust markerless augmented reality method for real-time applications by utilizing lines, circles, cylinders and spheres[7], which are sometimes rare in scenes. Klein and Drummond proposed a combination of head-mounted camera and rate gyroscopes that can yield estimation of head position at video field rate[12]. As this approach uses visual information when it is available, the alignment of computer-generated graphics onto display is accurate. Ballot et al. showed that initial camera registration can be done by capturing calibration marks, and most of them are preexisting features such as corners of doors or walls[1].

Simon and Berger proposed a markerless camera registration method that can be applied both for indoor and outdoor scenes. As they utilize multiple planner structures and buildings usually include some planes, their approach can be applied to align building models and videos of surveillance cameras with HPC camera image[26].

Satoh et al. introduced the idea of using external cameras for accurate position and orientation estimation of user camera[23]. However, as a marker on a user should be recognized by a bird-view camera, it is not suitable for wide outdoor areas.

Piekarski et al. proposed a hybrid MR system that can be used both at indoor and outdoor scenes[20]. Although they use markers in indoor scenes, they do not use visual clues in outdoor scenes hence the accuracy of MR is degraded.

For outdoor scenes, advanced hybrid camera tracking methods that utilize geometric information of buildings have been proposed[22][11][3]. As these systems have pre-defined building models, they can predict image features that are coming into camera image when the camera moves. However, if the textures of the buildings are changed from the ones that are used to make building models, there may be a case where this type of approaches could not handle the visually-shifted image clues even though robust KLT tracker[29] is used.

We set four surveillance cameras tentatively and one HPC user in our experiment field. As we install more surveillance cameras and expect multiple users simultaneously in the field, we need to design multi-sensor integrated system like the one Brown et al. proposed for collaborative mobile augmented reality[5].

As for visualization techniques of outdoor scenes, VESPER of Spann and Kaufman[27] and AVE of Sebe at el.[24] provide integrated visualization environments for wide areas[27], but they do not discuss the visualization of occluding situations.

3. See-Through Function

See-through display is realized by utilizing surveillance cameras embedded in real world. By calibrating surveillance cameras beforehand, objects in a hidden area can be shown on a HPC display once after HPC camera is calibrated.

Pose estimation of HPC camera without artificial markers in an outdoor scene is essential in order to realize mixed-reality visualization because it is impractical to set specially designed calibration markers such as checker boards in outdoor scenes.

Instead of embedding artificial markers, we propose to use substructures of buildings as calibration markers. We assume that the shapes of the substructures can be obtained from CAD data in this paper. We call the substructures landmarks.

One of the critical problems of this approach is that a landmark may look different at HPC camera even when the shape of the landmark can be precisely predicted. It may derive from light condition changes, color fading, aging, etc. These phenomena can not be inevitable even if relatively big and substantial parts of the buildings in a scene are set as landmarks. Therefore, we propose to extract live textures of the landmarks from video images of surveillance cameras every time when they are referred.

Figure 1 show the overview of our mixed reality system.

Three types of visual information are integrated into one image and it is shown to a user on HPC display. The α blending value of the live video is weighted larger than the other ones because it is the most important. On the contrary, the overlapped HPC camera image is assigned to the least blending value because a user can see the real scene with his/her own eyes just by looking

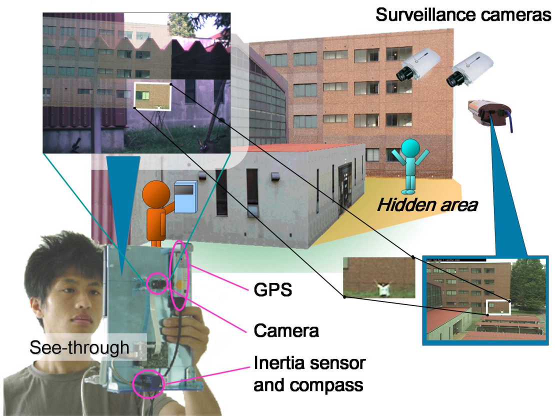

Figure 1. Overview.

away from the HPC display. We set 0.6 for the blending value of the live video, 0.28 for the CG models, and 0.12 for the HPC camera image. Figure 2 shows how video images and CAD models are integrated to show invisible areas to users. As the HPC camera is set just behind the HPC display, the user can recognize what is happening in the invisible areas as if he/she looks through the obstacles, e.g. buildings, walls, etc.

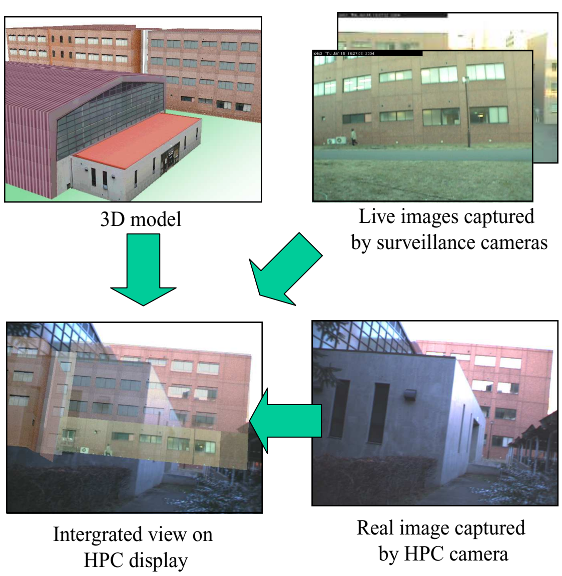

Figure 2. Integrated view on HPC display.

4. Camera Registration

Our HPC is equipped with a GPS, a digital compass an inertia sensor, and a CCD camera. GPS and digital compass are used to obtain initial estimation of location and orientation of the HPC . The inertia sensor can track the pose change of the HPC at high frequency. However, it is not sufficient to superimpose 3D models and live videos onto HPC camera image precisely because of its long-term instability.

For example, Figure 3(upper) shows an example of failed overlay with only GPS, compass, and inertia sensor. This snapshot is taken at the position A in Figure 11. The building is more than 20 meters away.

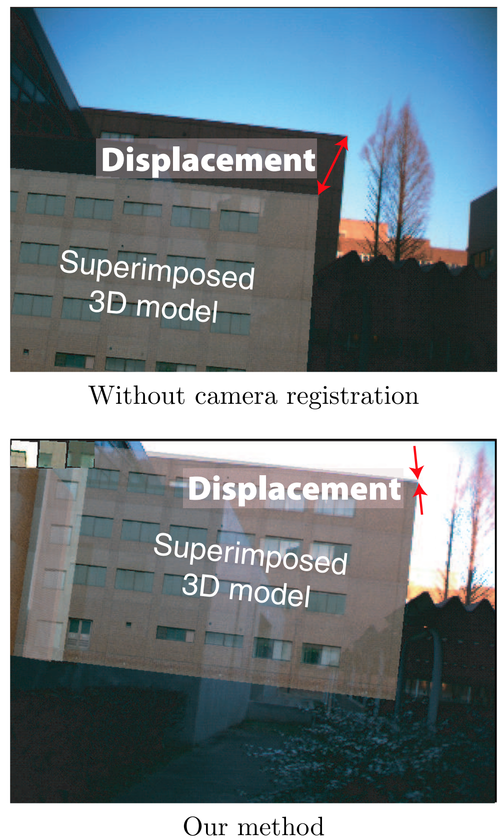

Figure 3. Example of camera registration.

Therefore, we define landmarks in the outdoor scenes and use them as calibration markers. As the landmarks are searched by image processing, they should be distinct in the scene. We adopt substructures of buildings that have good features to track[25] when they are projected on the HPC camera image. Choosing these substructures contributes to reduce the chance of false extraction in image processing.

We define two kinds of landmarks. **Primary landmarks** are the substructures that will be observed at skylines of buildings and have some vertical components. A typical example of primary landmarks is a corner of square buildings. One primary landmark is

0-7695-2191-6/04 $20.00 © 2004 IEEE

associated with several **secondary landmarks** that are set on the visible surface of the building. They should have a chance to be observed when the primary landmark is visible from a certain surveillance camera. Secondary landmarks are also asked to have horizontal and vertical features to avoid false matching. Figure 4 shows a set of one primary landmark and three secondary landmarks.

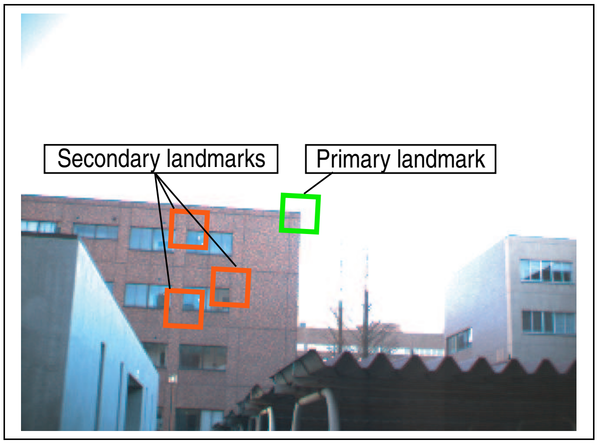

Figure 4. Primary and secondary landmarks.

Each landmark has geometric information given by the CAD data and pictorial information obtained by a surveillance camera in real time. The surveillance camera that takes pictorial information of the landmark is selected according to the locations of the HPC camera, the landmark, and the surveillance camera (Figure 5). When a landmark is to be searched on the image plane of a HPC camera, the corresponding image segment is warped by an appropriate projection matrix like the method of the real-time affine region tracker[9]. We call the warped segment a **landmark template**.

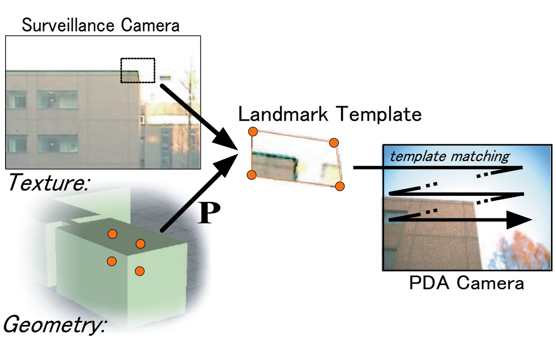

Figure 5. Geometric and pictorial information of a landmark.

The camera registration procedure is shown below.

1. Skyline area in a HPC camera image is segmented.
2. Distinct regions detected inside the skyline area by image processing[25] are selected as primary landmark candidates.
3. Among all the primary landmarks in the CAD model database, visible primary landmarks are selected based on the current status of pose estimation given by non-vision sensors.
4. Each pair of a visible primary landmark and a primary landmark candidate is examined by estimating the distance between the associated secondary landmarks and corresponding predicted positions on the HPC image. Prediction is calculated by fitting the visible primary landmark to the primary landmark candidate. If sufficient number of secondary landmarks are found close enough, the orientation and location of the HPC camera are recalculated based on the pair by applying ICP algorithm[4][30].

Figure 6 shows an example of initial edge detection of a HPC camera image. The upper figure of Figure 7 illustrates the skyline detection process and the lower figure shows the detected skyline area. Note that the skyline is not necessarily connected, because our objective is to estimate the positions of primary landmark candidates that are a set of points. The black dots in Figure 8 indicate the primary landmark candidates. As the landmark candidates should have orthogonally distributed differential components, most of the skyline part will be rejected.

5. Visualizing Invisible Area

In order to visualize objects in invisible areas, we use simple rectangular based video warping method in this paper.

Rectangles are set perpendicular to the ground in the invisible areas. Corresponding video segments are transmitted from an appropriate surveillance camera that can film the area, and the segments are warped based on the estimated projection matrices.

Figure 9 shows an example of the visualization. In this case, a rectangle is set on a wall of the building that is adjacent to the invisible area. The system holds the geometric information of the rectangle $(g_0, \cdots, g_3)$ and the corresponding live video is taken at $(c_0, \cdots, c_3)$ on the image plane of the most appropriate surveillance camera.

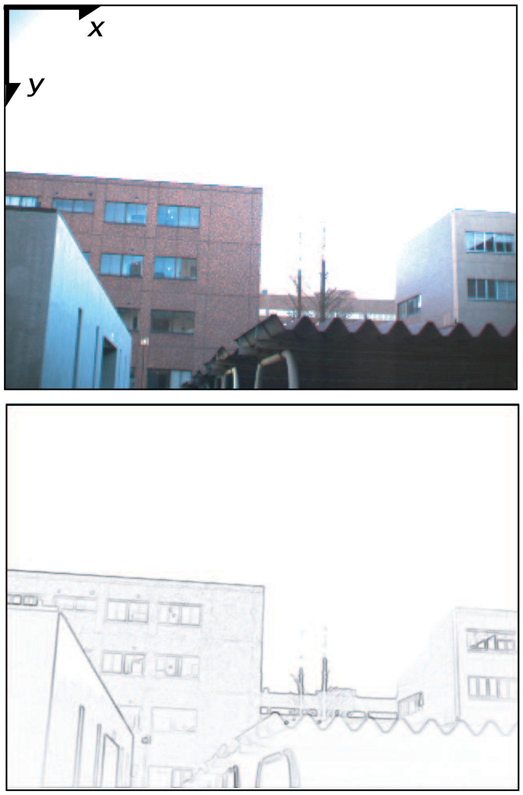

Figure 6. Edge detection.

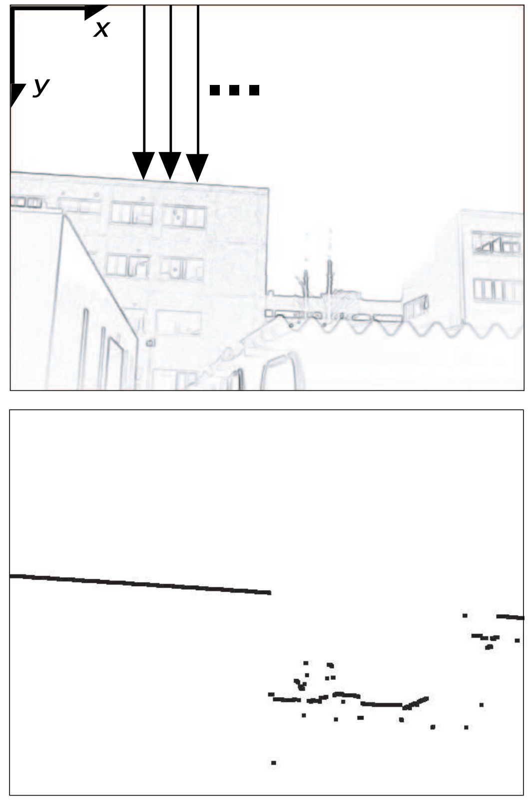

Figure 7. Skyline detection.

Figure 8. Primary landmark candidates.

This approach is simple but it has some limitations on locating the hidden objects accurately on the display especially when the objects are not so close to the rectangles. To improve the accuracy, we are planning to apply our billboard methods[15][10] that we have proposed in other papers to visualize the objects in the invisible areas. A similar visualization is also proposed in [24].

6. Experiments

We have implemented a preliminary system on the campus of University of Tsukuba. NEC Lavie LJ700/7E with Pentium-M 1GHz is used as the base of the HPC. The HPC is equipped with SONY IPS-8000 GPS sensor, a dragonfly IEEE1394 camera of Point Grey Research, and InertiaCube^2 of INTERSENSE. We use four Axis2120 web cameras as surveillance cameras. As the experiment environment is not permanent currently, the surveillance cameras are set on tripods at the locations shown in Figure 11. A user holds the HPC by hand when expriments are conducted.

Figure 10 shows the 2D map of our experiment field. An overview of the experiment field is shown in Figure 11. This photo is taken from the rooftop at point R in Figure 10.

When a user wants to improve geometric registration of mixed reality on HPC display, he/she directs the HPC camera to one of the buildings near his/her location, frames its skyline at the HPC camera, and requests calibration process. (Before: Figure 3 (upper). After: Figure 3 (lower).) Once calibrated, it is not nec-

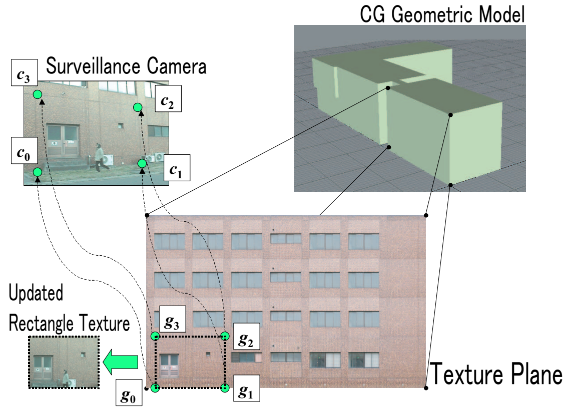

Figure 9. Visualizing invisible areas by updating textures. Video segments are sent from surveillance cameras.

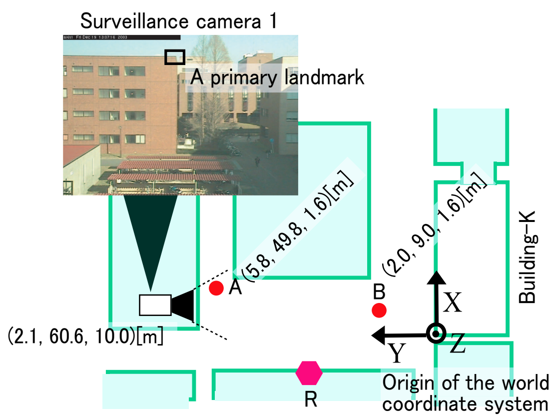

Figure 10. Experiment field.

essary to re-run calibration process until drift error of inertia sensor is accumulated. Figure 12 is a snapshot of HPC display overlaying a hidden area with a walker.

We have conducted an experiment to show the accuracy of our calibration process. Table 1 shows the results. A primary landmark used in this simulation experiment is a corner of the building-K (see Figure 10) and two locations A and B are also shown in the figure where A = (5.8, 49.8, 1.6) [meter] and B = (2.0, 9.0, 1.6) [meter]. At the both locations, HPC camera is set to direct the primary landmark roughly so that a surveillance camera 1 at (2.1, 60.6, 10.0) [meter] is associated for updating the landmark texture. The size of the landmark texture is 47 pixels wide and 31 pixels high. We took samples in the morning and in the evening to validate our method for daylight changes.

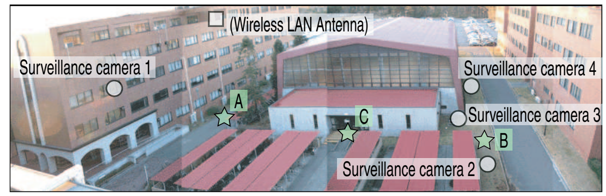

Figure 11. Experiment field overview.

Time in Table 1 indicates the time of taking samples.

For each sample, we measured the true location and orientation of the HPC camera beforehand. Then, certain offset is added to the true value and it is fed to the system as input. The offset values are shown in X, Y, Z, and *yaw* in Table 1.

The HPC camera captures images at the size of 640 pixels by 480 pixels. The location of the landmark corner on the HPC camera images are shown in *True* in Table 1 and the estimated location is shown in *Est*. The results showed that our approach succeeded in accurate camera registration.

The images captured in the experiment are shown in Figure 13. A box on the corner of the building (shown in bold red square) shows the estimated location of the landmark corner. A box that is apart from the corner (shown in thin green square) indicates the location of the landmark corner if the offset is given and no adjustment is conducted.

We also conducted another experiment for evaluating our calibration process. The calibration process was requested at the location C ten times by changing the direction of the HPC camera. In this experiment, the primary landmark of the building corner was used and two associated secondary landmarks were adopted. The displacement result between the true location of the primary landmark and the estimated one was 3.72 pixels in average. Total displacement result of the landmarks including the secondary landmarks was 10.55 pixels in average.

Our prototype system takes 450 ms in average to conduct the calibration process in one thread. This is invoked when a user requests to do so. Another thread is used to display overlapped HPC images with CG building models and warped videos of the surveillance cameras. This display thread runs at 15 fps with the projection matrix that is updated by the output of InertiaCube^2. Just after the calibration process is finished, the result is merged into the updating calculation of the projection matrix in the display thread.

In the experiments, surveillance cameras are also used to provide live textures of the building-K. The layout of the texture segments of the building-K is shown

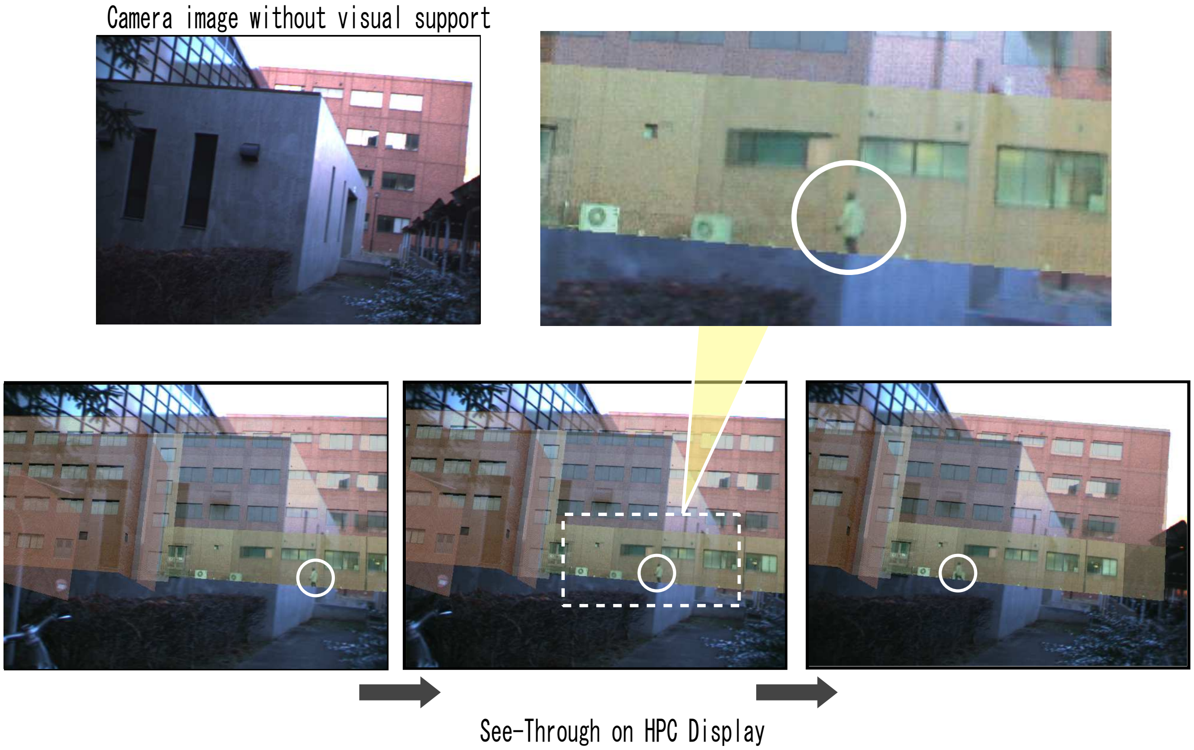

Figure 12. Real-time display of invisible areas.

in Figure 14. Three surveillance cameras are used to cover certain portions (α, β, and γ) of the building-K.

The final result of our mixed-reality system can be shown in Figure 12.

7. Conclusion

We proposed a new outdoor mixed-reality system with a camera-attached HPC in a scene where surveillance cameras are embedded. By our method, a HPC can display live status of invisible areas hidden by some structures such as buildings, walls, etc. The videos of the invisible areas are taken by surveillance cameras and they are precisely overlapped onto the video of the HPC camera.

We implemented a preliminary system with four surveillance cameras and the experimental results showed that our system can show invisible areas in real time.

References

[1] Y. Baillot, S. J. Julier, D. Brown, and M. A. Livingston. A tracker alignment framework for augmented reality. *International Symposium on Mixed and Augmented Reality*, pages 142–150, 2003.

[2] G. Baratoff, A. Neubeck, and H. Regenbrecht. Interactive multi-marker calibration for augmented reality applications. *International Symposium on Mixed and Augmented Reality*, pages 107–116, 2002.

[3] R. Behringer, J. Park, and V. Sundareswaran. Model-based visual tracking for outdoor augmented reality applications. *International Symposium on Mixed and Augmented Reality*, pages 277–278, 2002.

[4] P. J. Besl and N. D. McKay. A method for registration of 3-d shapes. *IEEE Transactions on PAMI*, 14(2):239–256, 1992.

[5] D. Brown, S. Julier, Y. Baillot, and M. A. Livingston. An event-based data distribution mechanism for collaborative mobile augmented reality and virtual environments. *Proceedings of IEEE Virtual Reality (VR2003)*, pages 23–29, 2003.

[6] K. W. Chia, A. D. Cheok, and S. J. Prince. Online 6dof augmented reality registration from natual features. *International Symposium on Mixed and Augmented Reality*, pages 305–313, 2002.

[7] A. I. Comport, Éric Marchand, and F. Chaumette. A real-time tracker for markerless augmented reality. *International Symposium on Mixed and Augmented Reality*, pages 36–45, 2003.

[8] Y. Cui and S. S. Ge. Autonomous vehicle positioning with gps in urban canyon environments. *IEEE*

Table 1. Matching result.

Sample	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Point	A	A	B	B	A	A	B	B
Time	9:20	9:21	9:31	9:32	16:11	16:13	16:21	16:24
X[m]	-10.0	10.0	-	-	-	-	-6.0	-
Y[m]	-10.0	10.0	-	-	-	-	-3.0	-
Z[m]	-	-	-	-	-	-	-	-4.0
Yaw[deg]	-	-	-21.0	21.0	-9.0	9.0	-	-
True	(307,241)	(322,240)	(358,246)	(339,354)	(342,224)	(325,225)	(318,252)	(365,247)
Est.	(307,239)	(323,239)	(356,243)	(341,355)	(342,222)	(325,222)	(317,250)	(365,248)

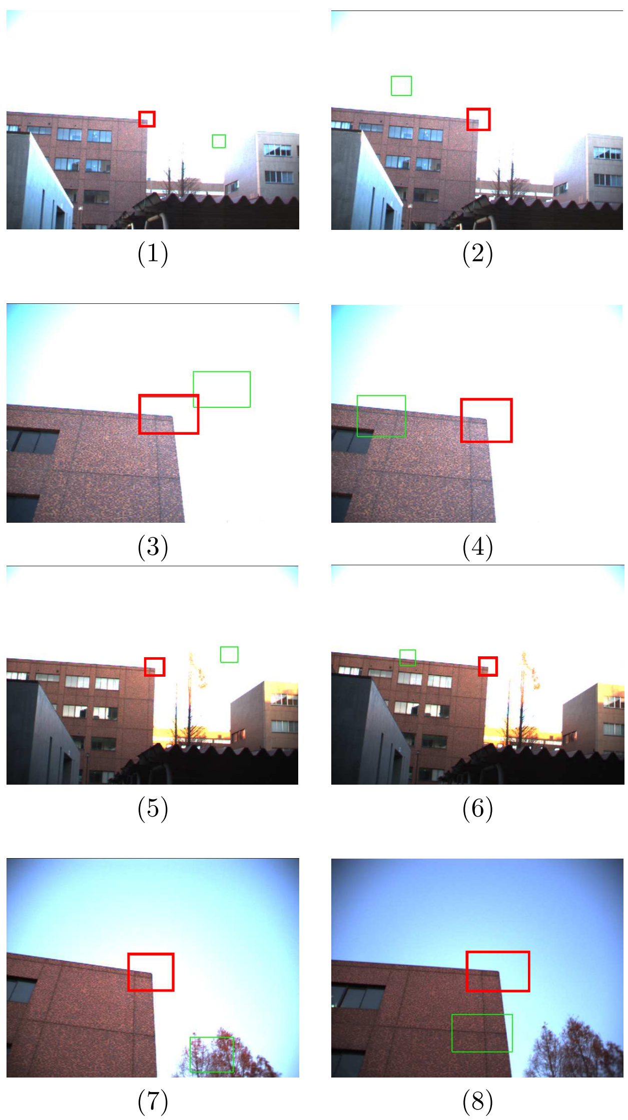

Figure 13. Images used for matching.

Transactions on Robotics and Automation, 19(1):15–25, 2003.

[9] V. Ferrari, T. Tuytelaars, and L. V. Gool. Markerless augmented reality with a real-time affine region tracker. *International Symposium on Argumented Reality*, pages 87–96, 2001.

[10] E. Ichihara, H. Takao, and Y. Ohta. Bird's-eye view for highway drivers using roadside cameras. *IEEE International Conference on Multimedia Computing and Systems*, 2:640–645, 1999.

[11] B. Jiang, U. Neumann, and S. You. A robust hybrid tracking system for outdoor augmented reality. *IEEE Virtual Reality*, pages 3–10, 2004.

[12] G. Klein and T. Drummond. Robust visual tracking for non-instrumented augumented reality. *International Symposium on Mixed and Augmented Reality*, pages 113–122, 2003.

[13] M. Kourogi and T. Kurata. A method of personal positioning based on sensor data fusion of wearable camera and self-contained sensors. *Proceedings of IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems (MFI2003)*, pages 287–292, 2003.

[14] M. Kourogi and T. Kurata. Personal positioning based on walking locomotion analysis with self-contained sensors and a wearable camera. *International Symposium on Mixed and Augmented Reality*, pages 103–112, 2003.

[15] T. Koyama, I. Kitahara, and Y. Ohta. Live mixed-reality 3d video in soccer stadium. *International Symposium on Mixed and Augmented Reality*, pages 178–187, 2003.

[16] J. Krösche, J. Baldzer, and S. Boll. Mobidenk - mobile multimedia in monument conservation. *IEEE Multimedia*, 11(2):72–77, April-June 2004.

[17] R. Kumar, H. S. Sawhney, Y. Guo, S. Hsu, and S. Samarasekara. 3d manipulation of motion imagery. *International Conference on Image Processing*, 1:17–20, 2000.

[18] W. Pasman and C. Woodward. Implementation of an augmented reality system on a pda. *International Symposium on Mixed and Augmented Reality*, pages 276–277, 2003.

0-7695-2191-6/04 $20.00

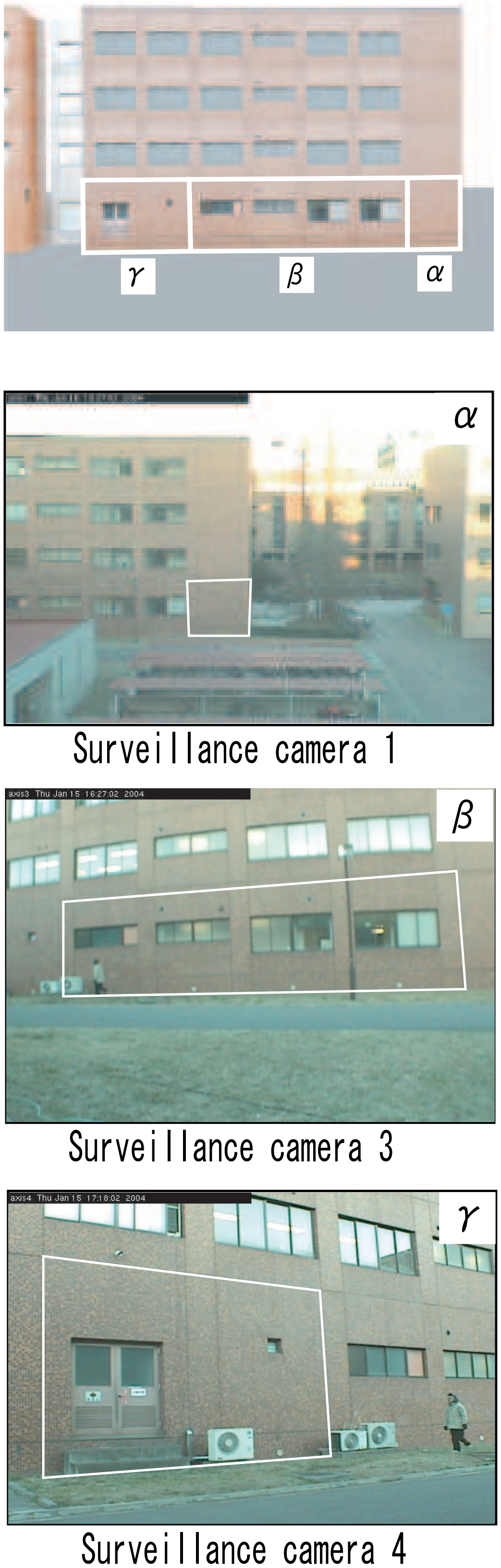

Figure 14. Generated textures from surveillance cameras.

[19] J. N. K. Patrick Dähne. Archeoguide: System architecture of a mobile outdoor augumented reality system. *International Symposium on Mixed and Augmented Reality*, pages 263–264, 2002.

[20] W. Piekarski and B. H. Thomas. Tinmith-metro: New outdoor techniques for creating city models with an augmented reality wearable computer. *International Symposium on Wearable Computers*, pages 31–38, 2001.

[21] N. B. Priyantha, A. Chakraborty, and H. Balakrishnan. The cricket location-support system. *Proc. 6th ACM International Conference on Mobile Computing and Networking (MOBICOM)*, pages 32–43, 2000.

[22] M. Ribo, P. Lang, H. Ganster, M. Brandner, C. Stock, and A. Pinz. Hybrid tracking for outdoor augmented reality applications. *IEEE Compuer Graphics and Applications*, pages 54–63, November/December 2002.

[23] K. Satoh, S. Uchiyama, H. Yamamoto, and H. Tamura. Robust vision-based registration utilizing bird's-eye view with user's view. *International Symposium on Mixed and Augmented Reality*, pages 46–55, 2003.

[24] I. O. Sebe, J. Hu, S. You, and U. Newmann. 3d video surveillance with augmented virtual environments. *ACM SIGMM Workshop on Video Serveillance*, pages 107–112, 2003.

[25] J. Shi and C. Tomasi. Good features to track. *IEEE International Conference on Computer Vision and Pattern Recognition (CVPR94)*, pages 593–600, 1994.

[26] G. Simon and M.-O. Berger. Reconstructing while registering: a novel approach for markerless augmented reality. *International Symposium on Mixed and Augmented Reality*, pages 285–294, 2002.

[27] J. R. Spann and K. S. Kaufman. Photogrammetry using 3d graphics and projective textures. *IAPRS*, 33(B5/1-V5):748–755, 2000.

[28] R. Tenmoku, M. Kanbara, and N. Yokoya. A wearable augmented reality system using positioning infrastructures and a pedometer. *Proceedings of IEEE International Symposium on Wearable Computers*, pages 110–117, 2003.

[29] C. Tomasi and T. Kanade. Detection and tracking of point features. *Carnegie Mellon University Technical Report CMU-CS-91-132*, April 1991.

[30] S. Uchiyama, K. Takemoto, K. Satoh, H. Yamamoto, and H. Tamura. Mr platform: A basic body on which mixed reality applications are built. *International Symposium on Mixed and Augmented Reality*, pages 246–253, 2003.

[31] D. Wagner and D. Schmalstieg. First steps towards handheld augmented reality. *International Symposium on Wearable Computers*, pages 127–135, 2003.

[32] X. Zhang, S. Fronz, and N. Navab. Visual marker detection and decoding in ar systems: A comparative study. *International Symposium on Mixed and Augmented Reality*, pages 97–106, 2002.