MR based Visualization of Viewing Fields of Surveillance Cameras in Outdoor Scene

Yoshinari Kameda   Dao Ngoc Thanh   Itaru Kitahara   Yuichi Ohta
Graduate School of Systems and Information Engineering
University of Tsukuba
{kameda,kitahara,ohta}@iit.tsukuba.ac.jp

Abstract: We propose a new MR based visualization method which can help people to find the viewing fields of surveillance cameras in outdoor scene. Though the surveillance cameras are always an active subject in our society, targeted people do not have even a tool to recognize where the monitoring areas are. Our proposed method can show the viewing fields in MR fashion, at the moment they need and at the place they want to check. We have implemented our preliminary system in which users exploit iPhone/VAIO-U for visualization and succeeded in showing the fields in five different ways.

1 INTRODUCTION

Surveillance cameras have become popular and placed almost everywhere, e.g. train stations, airports, banks, shops, streets. The balance between security and privacy is an active subject for political/social debates. Though it is not easy to keep a balance between security and privacy, we can say that a person may feel protected when he knows himself and his surroundings are being monitored by security cameras. Or, in some cases, one might feel uncomfortable for being monitored and want to find out the place where the cameras are not ready.

In either case, they probably want to know the viewing fields of surveillance cameras, e.g. where the cameras are, whether he/she is inside the observed area.

For indoor scenes, the space is usually small, hence it is relatively easy for people to locate the position of the cameras and to estimate their viewing fields.

On the contrary, In outdoor scenes, it is sometimes difficult to find out surveillance cameras that actually monitor people. Thus they need a help to find cameras and their monitoring areas.

Fortunately, there is a novel technology which can be applied to this problem. That is Mixed Reality (MR). MR is the encompassing of both Augmented Reality and Augmented Virtuality, merging the real world in which we are living and the virtual world created by computers. MR produces a new environment where real and virtual entities can co-exist and interact on line.

We propose to provide a new video image on which viewing fields of surveillance cameras are visualized by aligning the fields precisely to the real world in the video image. We fully utilize the MR technology to realize the system. Figure 1 illustrates our concept. A person is passing through a parking lot where a number of cameras are monitoring. He can easily be aware of monitoring areas in the scene with the help of MR based visualization tool.

With the proposed system, a person can check where the viewing fields are by looking at a mobile display. He can switch MR mode (on the bottom left picture) and map mode. In MR mode, five different visualization methods are prepared and he can choose one on-line, depending on his preference. Map mode, which also improves user's recognition of viewing fields, is also prepared.

As the available hardware in outdoor MR is limited due to the portability, we propose a system design in which a CPU consuming part is loaded to a remote CPU server.

Visual surveillance and monitoring (VSAM)[1] is one of the major topics in computer vision literature, but they do not discuss how the monitoring areas should be informed to targeted people.

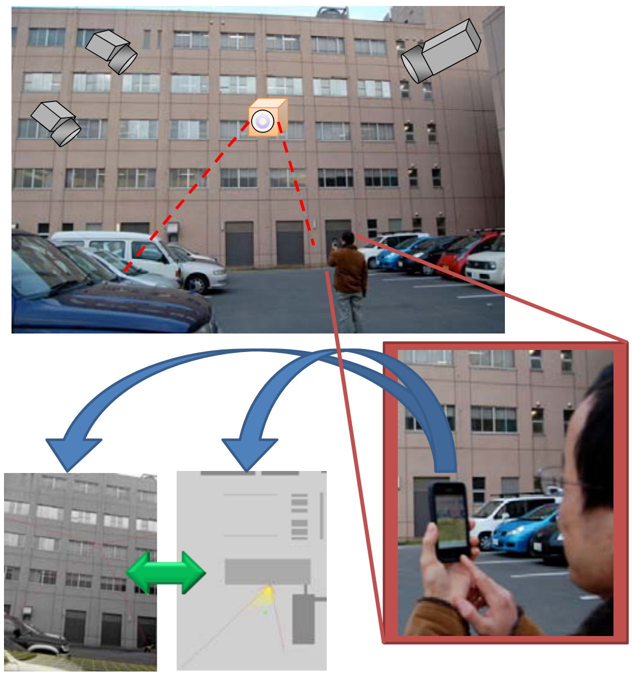


Figure 1: Visualization of viewing fields of surveillance cameras

2 SYSTEM DESIGN

2.1 Surveillance Camera

As new surveillance cameras are not frequently installed, we assume that the geometric property of surveillance cameras is given to the system in advance. The property should include the focal length of the camera or its viewing angle (field of view), the orientation of the camera, the position, and the resolution. They are registered in the world coordinate system so that monitoring areas could be figured out (Figure 2). We can also add any available properties to the camera description.

Figure 2: Safe area (monitoring area)

2.2 Portable device

In order to visualize the viewing fields of the cameras in user vision, a head mount display (HMD) is a straightforward solution on selecting a display device in MR literature. However, HMDs are usually bulky and inconvenient for outdoor applications, since high mobility and wide-range view are crucial for people in outdoor scenes. Therefore we need to design a more sophisticated MR system.

In recent years, mobile devices have become popular. Devices like cell phones and personal digital assistants (PDA) can be seen everywhere. Their prices have come down and they could reach the hands of ordinary people. As they can be equipped with some extra devices, e.g. GPS receiver and accelerometer, which are also useful for MR system, they are considered to be a good choice.

In outdoor MR, in addition to the display problem, we also take care of camera registration accuracy because it is critical for MR quality. As a consequence, the portable device should meet our demands which are listed below. For hardware:

1. Display: the larger display is better. However, it should not spoil the portability of the device itself.
2. Camera: it plays the role of user's view. It should continuously capture the image.
3. Network interface: as the portable device itself does not have sufficient processing power, the connection to a remote server should be established.

The following hardware functions are preferred to improve the accuracy of the camera registration. Note that we currently do not involve these functions to our preliminary system, but its integration could be done in near future.

4. GPS receiver: it is useful to estimate the location of a user. Unfortunately, the accuracy of GPS is not sufficient to our demand, image based camera registration method is still needed even if a GPS is available.
5. Accelerometer: if it is available, the system does not need to rely on computer vision all the time. However, the accelerometer which can estimate the orientation of the device very accurately is still expensive.
6. Compass: it is also useful if it can estimate the horizontal orientation of the mobile device / camera.

As for the software functionality, it should have:

A. Integrated rendering for mixing video image and 3DCG objects: it should run on line.
B. Camera registration: computer vision based approach is applied in this paper. Since the processing power is insufficient to run the registration on line, we make use of a remote CPU server. Detail should be discussed in the later section.

In this paper, we exploit ultra mobile PC(UMPC) and intelligent cell phone that meet our demands.

2.3 Computation Empowerment

Although computer vision based camera registration[2][3] is really helpful to get higher accuracy, it generally needs large amount of memory and powerful computing capability. Even current high-end cell phones or UMPC may not run MR applications properly. Therefore, we propose a new approach of computation empowerment for the cell phone and UMPCs (hereby we call PDA).

We introduce a remote CPU server which takes the role of image processing part. The mobile device sends a captured image to the remote server and receives the camera registration result in return. While the mobile device is communicating to the remote server, other processes including the rendering should run concurrently in order to avoid the performance down.

3 VISUALIZATION OF VIEWING FILED

Our goal is to provide a MR-based visualization method which makes people easily understand where the actual viewing fields of the surveillance cameras are.

One of the very simple solutions would be just showing the boundary of the viewing field on the ground surface. However, this solution is not sufficient to inform the actual viewing fields because it shows neither the direction the camera faces nor cubic expansion of the viewing fields.

On visualizing the viewing fields, the location and direction of the cameras also should be visualized in MR fashion so that people can intuitively understand them.

There are two new issues we need to take care of on defining visualization methods in MR fashion.

The first one is that inside-outside problem of viewing fields. The viewing fields are basically in pyramid shape, but people may sometimes see the pyramid from its inside, or from its outside. The visualization methods should be designed to handle both situations.

The second one is the overlapping problem. Overlapping may occur when the visualized pyramid is rendered, or when two or more pyramids are closely placed because there are multiple cameras in a scene. Face based rendering of pyramids may severely degrade the visibility of the actual scene which is overlapped by their surfaces, so wire based approach may be also useful in this sense.

The viewing field basically forms a pyramid shape of which bottom is cut by the ground. In case of some structures (e.g. buildings) are found within the viewing field, the pyramid should be cut by the structures. In order to estimate the partially cut viewing pyramid, geometric property of the buildings should be given in advance.

In this paper, we devise five different methods to visualize viewing fields of surveillance cameras.

(1) **“Edge”** method
Draw all the edges of the viewing pyramid. Shoulder edges and floor edges are rendered in different color. (In Figure 3(a), shoulder edges are in red and floor edges are in yellow.) The camera position is rendered as the closing point of the four shoulder edges. The rendering cost is very small, however, it is not easy to recognize the area of the viewing fields because users should find the square bounded by the floor edges

(2) **“Surface”** method
In addition to the edge, draw all side surfaces of the viewing pyramid translucently. This is useful when a user is not close to the viewing field. On the contrary, if the user is close to or in the viewing field, this visualization is useless because most of the view will be covered by the translucent surfaces.

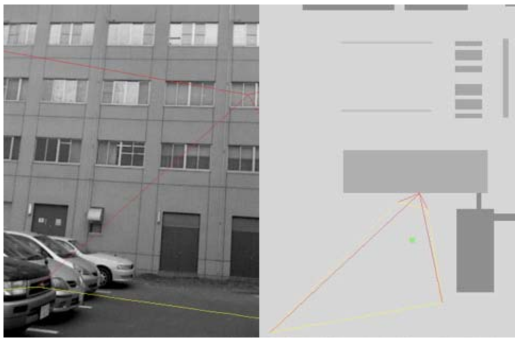
(3) **“Floor”** method
Similar to the surface method, but this approach draws the floor surface only. Translucent square drawn on the ground directly visualizes the lower end of a viewing field. This is useful especially when a user is close to / on the viewing field. If the user is far from the field, it is not so easy to estimate where the monitoring area is because the pyramid floor becomes very thin in his view. When a camera is close to the ground, the border region of the floor area could not cover the height of the people (if a person stands at the end of the pyramid floor area, his legs may be monitored, but the head could not be seen in video images).

(4) **“Arrow”** method
The floor area of the pyramid is sampled by a certain interval and an arrow is set at each sampling point, directing the camera position. The length of the arrow is inversely proportional to the distance to the camera position. A user can not only see where the viewing fields are but also estimate the direction of the camera and the distance to that just by looking at arrows. In Figure 3(d), blue dots denote the sample points on the floor surface.

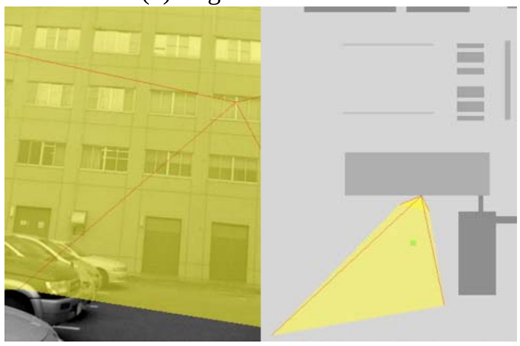
(5) **“Moving-floor”** method
The basic idea is similar to the “floor” method, but in this approach the floor of the pyramid is rendered as a grid-wise square, and the grid square is orthogonally set to the optical axis of a camera and moves along the axis. Since the four corners of the grid square runs on the shoulder edges, the grid square shrinks as it goes up to the focal point of a camera. A moving floor diminishes and disappears when it comes to the focal point of the camera.

Floor, Arrow and Moving-floor methods are designed mainly for people who stands closely or in viewing fields of surveillance cameras.

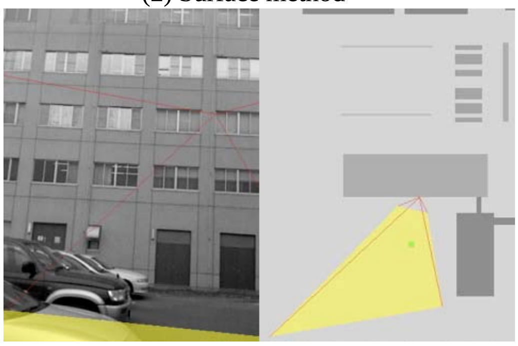
Left pictures of Figure 3 (a)-(e) show example shots of the proposed visualization methods that are implemented on i-phone. Note that users can switch any of visualization methods including map-mode (shown in the right pictures).



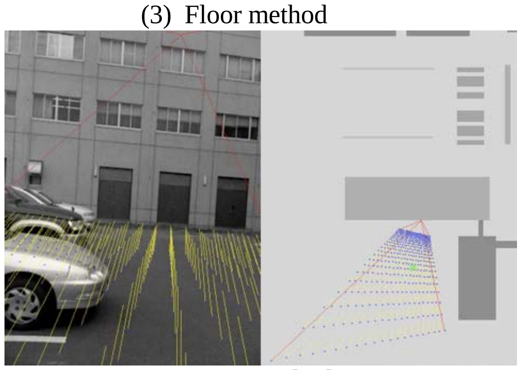
(1) Edge method



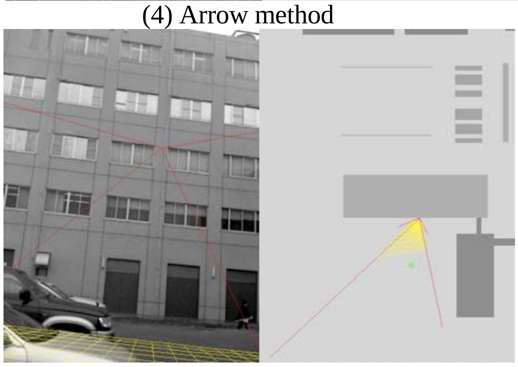
(2) Surface method



(3) Floor method



(4) Arrow method



(5) Moving-floor method

Figure 3: Visualization methods

In the map view, dark gray regions indicate buildings or some structures such as a parking lot of bicycles.

4 SERVER-SIDE PTAM

Accuracy and applicability of camera registration is crucial to archive high quality of viewing field visualization on live video images. We have selected Parallel tracking and mapping method (PTAM)[3] for camera calibration. PTAM utilizes natural features for reference points and estimate both the camera parameters (position and orientation) and 3D coordinates of the feature points simultaneously in high accuracy. However, as it is designed for dual core processor, it is difficult to run PTAM on mobile device.

We have introduced server-side PTAM to our preliminary system. The mobile device (client) continuously sends the green channel of the captured image in compressed format to a remote server. The remote server also continuously sends the estimated camera parameters to the client (Figure 4).

Since the original PTAM generates a 3D scene model based on the obtained feature points on every bootstrap, we need to invent a new method to mix up the surveillance camera properties (geometric information of camera viewing pyramid) to the 3D scene model. We have realized it by embedding the pyramid into the 3D scene model that is pre-built before the service starts. Instead of normal bootstrapping of PTAM, our system loads the pre-built 3D scene model and start searching the corresponding feature points in video images.

5 IMPLEMENTATION

A prototype system has been implemented and it successfully provides Mixed reality video that shows superimposed camera viewing fields onto its actual video image at a couple of frames per second with the help of the PTAM server on the network.

As for mobile device, we picked up Sony VAIO VGN-UX90PS and iPhone as PDA, and exploit an Apple MacBook Pro as a remore CPU server. This results in a prototype providing on-line video with no requirement for any GPS or gyrocompass devices. Figure 4 shows the design of our server-client architecture. The prototype system successfully runs in our campus (Figure 5) where we prepare 3D building models (Figure 6) in order to visualize the viewing fields of the surveillance cameras more precisely. A red rectangle in Figure 5 indicates the area shown in the map mode in Figure 3 (by iPhone).

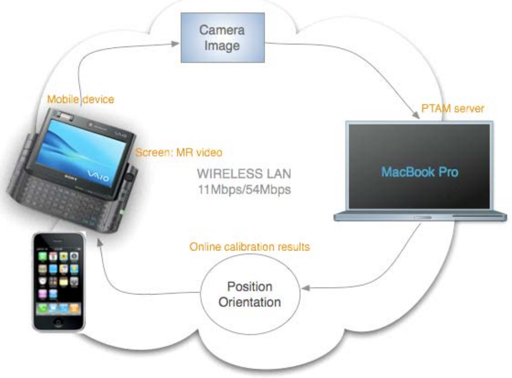


Figure 4: Server-side PTAM

Figure 5: Experiment area

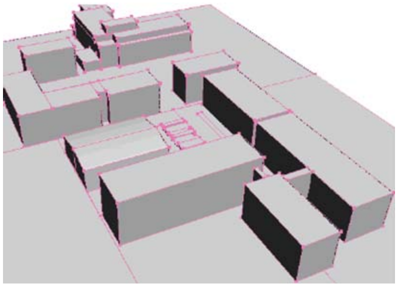


Figure 6: 3D models in experiment area

6 CONCLUSION

We propose a new MR based visualization method of viewing fields of surveillance cameras in outdoor scene. People can be aware of the monitoring and recognize the precise size and location of the viewing fields. As the viewing fields are rendered in the real-time video, they can easily estimate where the monitoring areas are.

Currently we have provided five different visualization methods (plus map view) with the choice of users. Further investigation will reveal the necessary functionalities on visualization of the viewing pyramids for better understanding. Our preliminary system updates the video image on the mobile device display at most 7 fps (by iPhone), hence we need to speed up the system and improve the accuracy. We may combine extra sensors to reach the higher goal.

REFERENCE

[1] Takeo Kanade, Robert Collins, Alan Lipton, P. Anandan, Peter Burt, and Lambert Wixson, "Cooperative multi-sensor video surveillance," DARPA Image Understanding Workshop, pp. 3-10, 1997.

[2] Gerhard Reitmayr and Tom Drummond, "Going out: Robust tracking for outdoor Augmented Reality," 5th Int. Symp. Mixed and Augumented Reality, pp.109–118, 2006.

[3] Georg Klein and David Murray, "Parallel tracking and mapping for small AR workspaces," 6th Int. Symp. Mixed and Augumented Reality, pp. 225–234, 2007.

[4] Masayuki Hayashi, Itaru Kitahara, Yoshinari Kameda, and Yuichi Ohta, "Client-server processing of PTAM for mobile Mixed Reality," Proc. IEICE General Conf. of Information and Systems Society, pp.163, 2009 (Japanese).