Smooth Video Hopping for Surveillance Cameras

Takahiro Tsuda* University of Tsukuba
Itaru Kitahara† University of Tsukuba
Yoshinari Kameda‡ University of Tsukuba
Yuichi Ohta§ University of Tsukuba

1 Introduction

As more and more surveillance cameras are installed in towns and cities, they are going to be more open to the public, and people may want to see these videos. In such a scenario, people may find it difficult to understand which place they are seeing because it is almost impossible to remember which camera images are from which place and from which direction. This problem is particularly important when a viewer changes from one camera to another to see a scene from a different angle. This happens because the viewer sometimes does not understand the spatial relationship among cameras in a real space. When the image is instantly switched from one camera to the other, the viewer may feel disoriented because he/she has to place the location of the new camera in his/her mind to recognize what he/she is seeing. Therefore, a sophisticated camera switching method is needed that enables viewers understand how camera viewpoints change.

We propose a new visualization application that allows viewers to "hop" from one camera to another. When a viewer is viewing one camera, the viewer sees only the video taken directly by the camera. Then, when the viewer wants to change the viewpoint to another camera, our system shows the viewer a pseudo-3D transition video sequence until the transition is complete. After the transition, the viewer can directly view video taken from the new position. As these transitions are like flying over a scene from one viewpoint to another, the viewer can easily understand the spatial relationship between the two cameras.

2 Transitions between Cameras

We created a pseudo-3D transition video of the hop from one camera to another using projection mapping of live video textures onto major static objects such as buildings as well as the ground in a scene.

We assume that the 3D shapes of all static objects are input to the system and that all cameras are calibrated in advance.

We include only the major objects and the ground for modeling the scene because our purpose is to assist the viewers to understand what they see when they change the viewpoint. We think that detailed scene observation can only be done when they view real video from a camera's viewpoint. Since people are not sensitive to shapes when they fly over a scene, it is not important to reconstruct a precise 3D world during the transition in this application.

The transition path between two cameras is set by linearly interpolating the rigid translation matrices of the cameras; the intrinsic parameter matrix is also calculated using linear interpolation.

The textures of the models are updated on-line. Each 3D model surface has at most two video sources.

If the surface is visible only from one camera, the video from that camera is mapped onto the surface throughout the transition using projection mapping.

*e-mail: tsuda@image.esys.tsukuba.ac.jp
†e-mail: kitahara@iit.tsukuba.ac.jp
‡e-mail: kameda@iit.tsukuba.ac.jp
§e-mail: ohta@acm.org

If the surface is visible by the both cameras, the videos of the two cameras are blended to create the surface texture. We propose sigmoidally changing the blending weight (from 0.0 to 1.0) during the transition. There are two reasons for sigmoidal blending. The first reason is that a mixed texture of around 0.5 is not good quality even when the cameras are well calibrated geometrically; this is because of color inconsistency. The second reason is that the viewers can perceive the switching of textures over the steep weight change, which means they can recognize the moment the cameras are switched.

3 Implementation

Figure 1 shows some snapshots of a pseudo-3D transition from camera-A to camera-B. The upper-left picture shows video taken by camera-A and the bottom-right one corresponds to video from camera-B. The transition from camera-A to camera-B is set to 3.0 seconds, and the textures in the scene are updated at more than 15.0 fps during the transition using SONY VAIO-U70 and two AXIS 2120 network cameras.

In this video transition, only two buildings are modeled in the scene. Although there are some small structures on the ground and they are displayed as distorted textures of the ground surface, they do not seriously affect the viewer's spatial perception during transition.

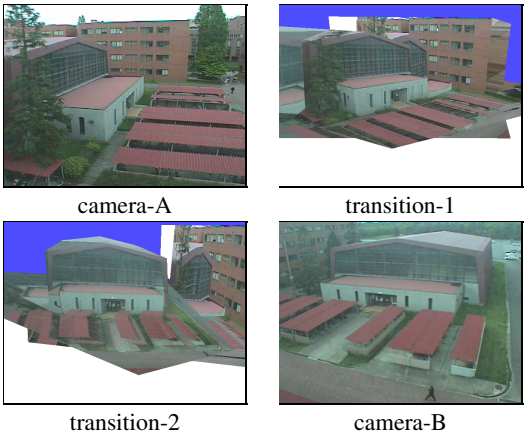

Figure 1: Pseudo-3D Transition

4 Conclusion

We have presented a new "video hopping" method for visually switching among surveillance cameras smoothly.

The system we proposed here is going to include a camera for a handheld PC(HPC) that is carried by a person in a scene. The HPC camera can be dynamically calibrated by GPS and an inertia sensor[2004]. Once they are integrated, a viewer watching the HPC camera image on a display can hop out of his/her standing point and shift to on one of the surveillance cameras by watching the smooth pseudo-3D transition.

[2004.] KAMEDA, Y., TAKEMASA, T., AND OHTA, Y. Outdoor mixed reality utilizing surveillance cameras. In *SIGGRAPH 2004 Sketches*.