Cooperative Distributed Vision for Dynamic Three Dimensional Scene Understanding
（ Project No. JSPS- RFTF 96P00501 ）

Project Leader
Takashi Matsuyama Professor, Graduate School of Informatics, Kyoto University
Core Member
Michihiko Minoh Professor, Center for Information and Multimedia, Kyoto University
Minoru Asada Professor, Faculty of Engineering, Osaka University
Toshikazu Wada Assistant Professor, Graduate School of Informatics, Kyoto University

1. Objective

The goal of the Cooperative Distributed Vision System is summarized as follows (Figure. 1):
Embed in the real world a group of network-connected Observation Stations (real time image processor with active camera(s)) and mobile robots with vision, and realize (1) real time dynamic real world scene understanding, and (2) versatile 3D scene visualization. We may call it Ubiquitous Vision.

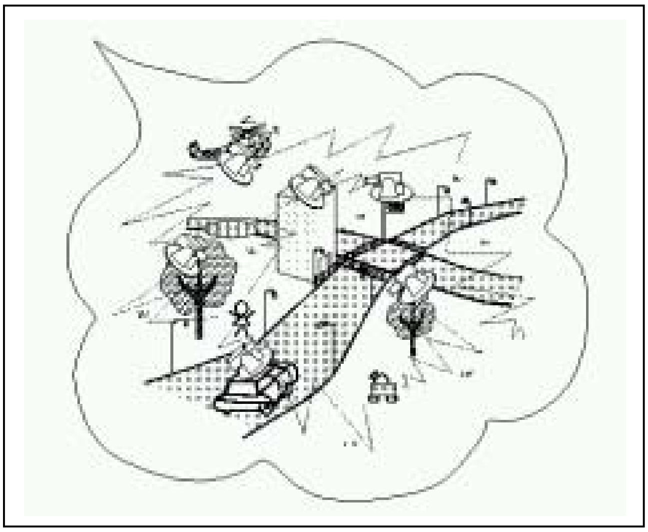

Figure 1: Cooperative distributed vision

2. Summary

2.1 Multi-Focus Camera for Real Time Depth Sensing

We developed a multi-focus camera that can capture three images of different focus values (i.e. different degrees of blur) at video rate: locations of three CCD planes in a 3CCD video camera are slightly shifted from normal positions to capture three blurred images simultaneously. We showed that the introduction of a coded aperture (i.e. iris with complex spatial pattern) greatly increases the accuracy of depth sensing (Figures 2 and 3).

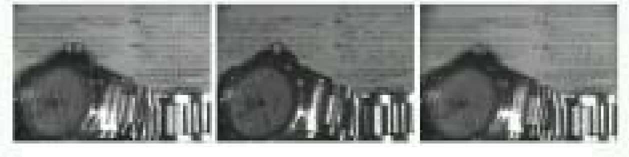

Figure 2 Three images of different degrees of blur captured by the multi-focus camera with a coded aperture.

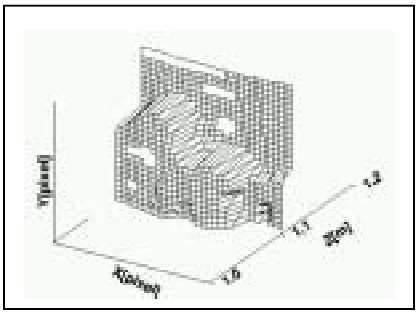

Figure 3 3D depth image of the watch computed from the multi-focus images in Figure 2.

2.2 Fixed Viewpoint Pan-Tilt-Zoom Camera

We developed the fixed-viewpoint pan-tilt-zoom (FV-PTZ) camera: an active camera whose optical center (viewpoint) is aligned at the rotation center and stays fixed irrelevant to the camera rotation. The FV-PTZ camera enables us to capture a very high resolution omni-directional panoramic image and moreover, facilitates the development of real time moving object tracking systems.

Figure 4 High resolution omni-directional panoramic image captured by the FV-PTZ camera.

2.3 Real Time Multi-Target Tracking

We developed a PC cluster, in which a group of PCs with FV-PTZ cameras are connected by an ultra high speed network (Figure 5). A multiple moving object tracking system was developed using this PC cluster system (Figure 6).

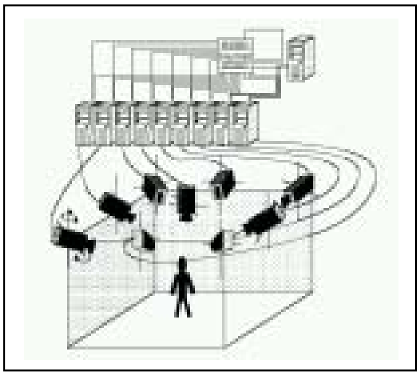

Figure 5 PC cluster with FV-PTZ cameras.

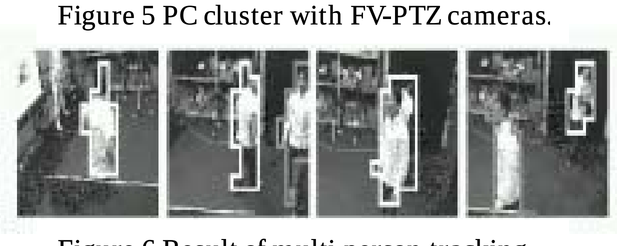

Figure 6 Result of multi-person tracking

2.4 3D Video Generation, Visualization and Editing

With the PC cluster system, a set of synchronized multi-viewpoint object motion images can be observed (Figure 7). Then, (1) extract multi-viewpoint object silhouettes from the observed images, (2) compute 3D object shape based on the silhouettes, (3) represent 3D object surface by a group of small planar patches, and (4) obtain texture (i.e. color pattern) on each patch from the observed images. (5) With these processes, a real 3D object image (3D Video) can be generated.

3D video is NOT artificial CG animation but an ultimate image medium recording real world object behaviors as is. Its applications include (1) 3D digital archive for traditional dances such as “noh” and Olympic games (2) 3D video encyclopedia for animal life documentary (3) rehabilitation for the disabled and training for golf playing (4) tele-presence communication system over networks, and so on.

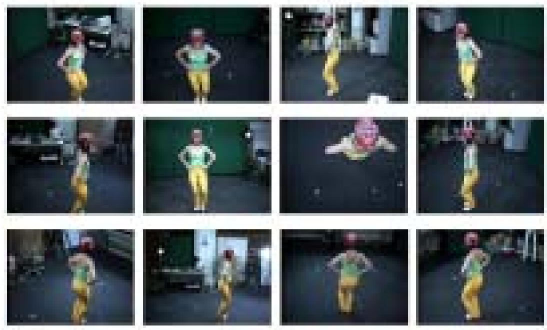

Figure 7 Multi-viewpoint object images

Figure 8 3D video composed with the omni-directional image in Figure 4.

3. Conclusion

The vision sensors, image analysis methods, and tracking systems developed by our project will greatly facilitate the development of various security monitoring and ITS systems. Moreover, 3D video will cut the new frontier of image media applications over digital TV and broad-band networks in the 21 century.

Major Papers

T.Matsuyama : Cooperative Distributed Vision, Lecture Notes in Artificial Intelligence, 1701, pp.75-88, 1999

T.Matsuyama and T.Takai: Generation, Visualization, and Editing of 3D Video, Proc. Of 3D Data Processing, Visualization, and Transmission, Padova, Italy, 2002