Cooperative Monitoring by Fixed Cameras and Moving Cameras

Sadaaki Kunimatsu, Hironori Kumeno, Naoko Tsukuda, Tomohito Takubo, Kenji Inoue, Tatsuo Arai
Graduate School of Engineering Science, Osaka University
1-3 Machikaneyama, Toyonaka, Osaka, 560-8531, JAPAN
kumeno@arai-lab.sys.es.osaka-u.ac.jp

Yasushi Mae
Faculty of Engineering
University of Fukui
3-9-1 Bunkyou, Fukui, Fukui, 910-8507, JAPAN

Tomohiro Umetani
School of Design and Architecture
Nagoya City University
2-1-10 Kitachikusa, Chikusa, Nagoya, Aichi, 464-0083, JAPAN

Abstract: In this paper we propose a human tracking system for dynamic environments based on the cooperation of fixed and moving cameras to create safe and secure spaces at construction sites. It is difficult for only the fixed cameras to monitor the all of a crowded dynamic environment because of the occlusion by the large objects such as the construction materials. We introduce moving cameras such as the mobile robots for monitoring these areas unobservable by fixed cameras. The moving cameras monitor the spaces unobservable of the fixed cameras that are caused by occlusion by the objects. The system is able to monitor a target position through cooperation of the fixed and moving cameras. Experimental results illustrate the effectiveness and feasibility of our proposed cooperative system.

Keywords: cooperative system, human tracking, safety and security, mobile robot

1. Introduction

The first priority is safety of workers at construction sites. Monitoring the states of workers and workspaces at construction sites in detail is one of the crucial issues for building a safe and secure construction site.

In order to expand the monitoring areas by fixed cameras, using cooperative monitoring systems based on multiple cameras is a reasonable solution for efficient monitoring. Several cooperative monitoring systems based on the fixed cameras have been developed [1],[2]. At construction sites, there are many dangerous and restricted areas. It is important to monitor the time when people enter such areas. The state of a construction site and the area unobservable of the fixed cameras located in it changes by the achievement of the construction tasks and the layout of the construction materials at the site. In addition, some workspaces of the construction site move according to the progress of the construction task. Therefore, it is difficult that the whole area is observed by only the fixed cameras, and it is desired that the viewpoints to monitor the workspace of the construction site should be changed according to the condition of it. Under such conditions, if we can completely observe the attractive area, we can contribute to a creation of safe and secure spaces.

In this research, we propose a method for cooperative monitoring systems based on moving cameras and fixed cameras at construction sites to create safe environments for workers. The fixed cameras monitor the assigned areas, respectively, and the moving cameras can move with some freedom and monitor the whole area. We introduce moving cameras such as mobile robots in the monitoring area for several reasons: to monitor the areas unobservable of the fixed cameras, to adjust the change of environment and to acquire the data of targeted persons in detail. The change of viewpoints of the moving cameras enables the monitoring system to achieve the above three functions.

One of the applications of the system is cooperation with data collection systems, for example, data logging systems and RFID devices. The cooperation of the proposed system and the data collection system yields recording the report of the construction task, the task analysis, and tracking of the construction materials at the construction site.

This paper introduces the cooperative system that consists of the fixed cameras, which are fixed in the environment, and the moving cameras, which can move in the environment in order to change their viewpoints to monitor all of environment. In dynamic environment, it is important to achieve effective monitoring motions by cooperation of fixed and moving cameras. If they monitor the environment independently, the cooperative system cannot monitor the targeted person continuously. For example, if there is not a cooperation rule for tracking the person, when the tracking person comes in the monitoring area of the fixed camera again, it is difficult that the fixed camera can distinguish whether a newly monitored person is the person whom the fixed camera have monitored or not. In addition, when the person goes out of the monitoring area of the camera and comes in the area of other camera, the tracking result is not integrated by the monitoring system. In this case, the continuous monitoring ends in failure.

We also introduce a method for detection of the targeted person for moving camera in the cooperative system using distance data and color data of the person. The moving camera is desired that the camera should detect and monitor the targeted person to realize the cooperative monitoring

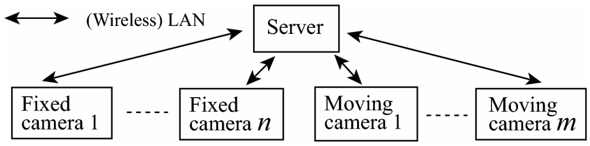

Fig.1. Relationship between server, fixed and moving cameras

system even if that camera moves. It is difficult to detect the person by the frame subtraction method when the camera is moving. We use the distance data and the color data, those data are robust against change of environmental conditions.

In this paper, we first state the cooperative system that we consider in this research. Then we describe methods of monitoring human behavior by using fixed and moving cameras. We have carried out the continuous person tracking experiment under each fixed camera has its blind spot. Finally, we illustrate our experimental environment and describe the experimental results to show the feasibility of our proposed method.

2. COOPERATIVE SYSTEM

This section describes the cooperative system. We introduce the concept of the system and its elements in this research, namely a server, fixed and moving cameras.

2.1 Cooperative System

Recently, cooperative systems have been proposed in the computer vision field [3],[4],[5] and the studies on human tracking by using mobile cameras [6],[7],[8],[9],[10],[11]. The cooperative system defined in this research is a system which consists of one server and many cameras connected by LAN(Local Area Network) as shown in Fig.1 and monitors human behavior by using them. As a cooperative system, we consider a system consisting of subsystems, where a subsystem has several cameras set in the assigned area and a mobile camera moving in it. The concepts of the system and the subsystem is shown in Fig.2 and 3, respectively, where Sub_{ij} denotes ij -th subsystem in Fig.2.

The moving stereo cameras are put in the space where there are several stereo cameras set in the fixed locations. The fixed cameras have their camera coordinate systems and can obtain the camera coordinates of objects, respectively. By transformation from the camera coordinates into the world coordinates, the objects can be treated as the coordinates of the identical space. Based on the condition that the moving cameras can obtain online the self positions, the camera coordinates of the moving cameras can also be transformed into the world coordinates.

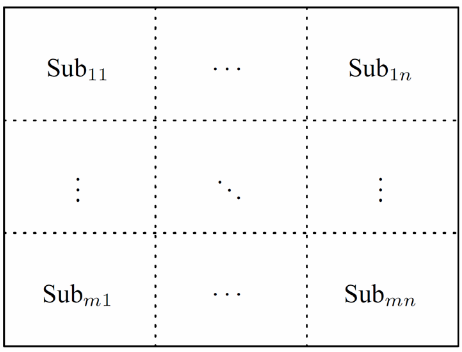

Fig.2. Concept of cooperative system

2.2 Server

The server in our system has the following three roles as a function:

1) Database
2) Supervisor
3) Visualization

Role 1) means that the database integrates the data sent from all the cameras to the server as a monitoring result and stores it. Role 2) means that the supervisor gives some orders to all the relevant cameras. Role 3) means that the visualization of important information such as the location of the targeted person is realized.

2.3 Fixed Cameras

Fixed cameras are set in the specified locations. The shared area between each fixed camera is made at a minimum, where the field of view of cameras in this paper is defined as the range that can observe until height 1.8 [m] from the ground. The fixed cameras sequentially send information to the server.

2.4 Moving Camera

Moving cameras are arranged at a rate of one per several, and are also connected to the server by wireless LAN, respectively. They can obtain color images and simultaneously measure distance from themselves to objects. We assume that the moving cameras run within the determined space, for example, on the rails. Note that the moving cameras need not necessarily be able to move everywhere in the assigned areas although they need to be able to monitor the objects almost everywhere. The moving cameras have the following two action modes.

1) Tracking mode
2) Standby mode

In the tracking mode, moving cameras track the targeted persons. In the standby mode, moving cameras act as fixed cameras of unobservable areas.

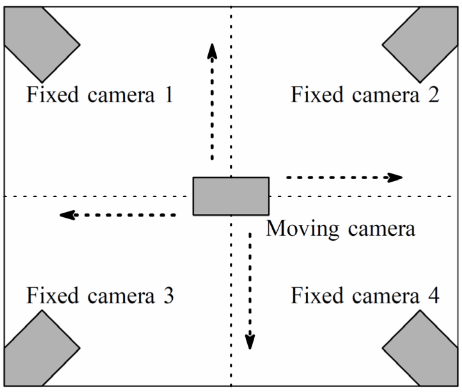

Fig.3. Concept of subsystem

1) Tracking Mode: Once the supervisor gives all the relevant cameras an order for monitoring the targeted person, the corresponding cameras start monitoring it based on the information such as the world coordinate obtained from the database server. Particularly, if the targeted person is going to an area unobservable of the fixed camera, the supervisor orders that the moving camera inherits the monitoring of the targeted person from the fixed one. If the supervisor sends a standby instruction to the corresponding moving camera, it transfers from the tracking mode to the standby mode.

2) Standby Mode: In the standby mode, a moving camera behaves as a fixed cameras of the nearest unobservable area.

3. MONITORING

In this section, we state detection methods of persons for monitoring human behavior. We use different two methods in fixed and moving cameras in order to detect and tracking persons. We discuss in this section the characteristics of the detection method for the moving camera using the detection result by the camera that is moving.

3.1 Monitoring by Fixed Cameras

In monitoring by fixed cameras, we use the background subtraction method. We detect a head of a person in order to monitor human behavior. We use a silhouette of a head to detect a head region. The head silhouette template $T(j,i)$ is represented by the following binarized image.

$$T(j,i) = \begin{cases} 1 & \text{if}(j,i) \text{ is on a silhouette} \\ 0 & \text{otherwise} \end{cases}$$

where (j,i) indicates the pixel coordinate on the template. We detect the corresponding head region from background subtraction images by matching to the head silhouette template, where we use SAD (Sum of Absolute Distances) as a criterion to detect the head region. The image region with the smallest SAD value that is lower than some predetermined threshold is detected as the head region.

3.2 Monitoring by Moving Cameras

In monitoring by moving cameras, it is difficult to apply the subtraction image to detect the targeted person like a case of fixed cameras. Therefore, we proposed the method based on the two information obtained by both color images and distance of the object to detect the targeted person. The procedure of detection of the targeted person is as follows:

1) Binarization Based on Distance: In this research, we can calculate distance of objects since also moving cameras are stereo cameras. Based on distance from the moving camera to the targeted person, we binarize the image obtained by the moving cameras. We specify its threshold as follows:

$$D_b(u,v) = \begin{cases} 1 & \text{if } d_l \le D(u,v) \le d_u \\ 0 & \text{otherwise} \end{cases}$$

where (u,v) indicates the pixel coordinate on the image, $D_b(u,v)$ is a binarized image, $D(u,v)$ is an image with distance values, d_l is the lower threshold, and d_u is the upper threshold. Each threshold is determined by the position of the targeted person at previous frame.

Through the process, the camera obtains the present person's position with respect to the camera coordination frame. The obtained person's position is used for the detection of the target position. In addition, the position is also used for setting the search area of the image frame to improve accuracy of the detected person's position and to reduce the calculation cost for detection method.

2) Binarization Based on Color: We use color information in order to detect robustly a person when the moving camera is running. In this paper, we distinguish the head of a person from background by detecting both colors of face and hair.

$$C_b(u,v) = \begin{cases} 1 & \text{if } |C_{rg}(u,v)| \le c_{bk},\ |C_{gb}(u,v)| \le c_{bk}, \\ & \quad |C_{br}(u,v)| \le c_{bk} \text{ and } h_l \le H(u,v) \le h_u \\ 0 & \text{otherwise} \end{cases}$$

where $C_{rg}(u,v)$ is a difference image between images of red and green in the RGB color space, $C_{gb}(u,v)$ is that between green and blue and $C_{br}(u,v)$ is that between blue and red. And c_{bk} is the threshold of each difference image, which is determined c_{bk} = 10, experimentally. $H(u,v)$ is an image of the Hue value in the HSL color space, h_l is the lower threshold and h_u is the upper threshold. The HSL color space has several advantages compared to the RGB color space, for example, its three elements have low correlativity, and Hue value is useful to extract the skin color, and it is hard to be influenced by change of lightness caused by the lighting condition.

3) Detection of the Head: We evaluate SAD with the following matching process.

Case of $T(j,i)=1$

If the corresponding $D_b(u,v)=1$ or $C_b(u,v)=1$, then the pixel (j,i) of the template and the corresponding pixel (j,i) of the image obtained by the moving camera are matching. Otherwise, not matching.

Case of $T(j,i)=1$

If the corresponding $D_b(u,v)=0$ and $C_b(u,v)=0$, then the pixel (j,i) of the template and the corresponding pixel (j,i) of the image obtained by the moving camera are matching. Otherwise, not matching.

This matching process means that distinction is loose inside the human silhouette of the template, while is strict outside that. By this method, we can achieve human tracking when the moving camera is moving.

Fig.4. Human tracking with pan motion by proposed method

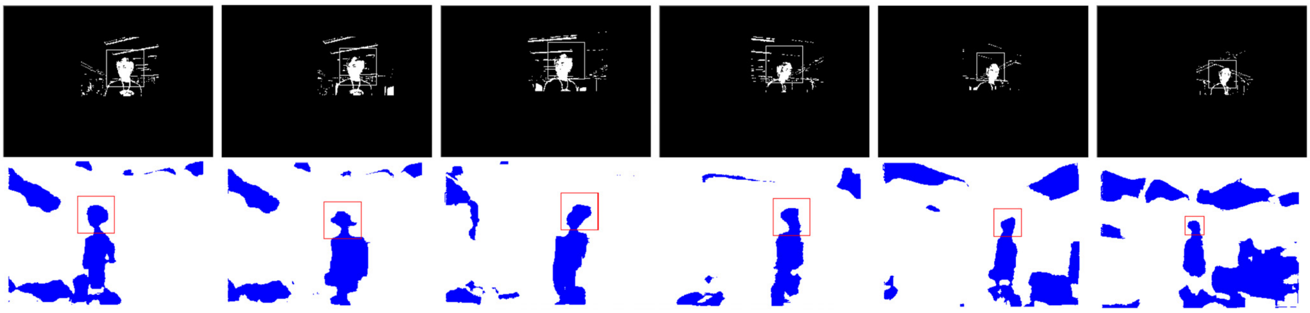

Fig.5. Human tracking with pan motion by proposed method, detection result by color data (upper), detection result by distance data (lower)

Fig.6. Human tracking with pan motion by frame subtraction method

3.3 Characteristics of Detection Method for Moving Cameras

We discuss the characteristics of the detection method of the moving camera through the detection result of the person by the moving camera. We show the detection results by the frame subtraction, those by only the distance, and those by the combination of the color information, respectively, while the camera is moving. We use a trinocular camera unit “Digiclops Stereo Vision” as the camera part of the moving camera. The whole image size of the camera is set 320×240 [pixels]. The camera unit is rotated (pan motion), and the camera unit is not given the motion information beforehand.

We evaluate the detection method based on the sequences of the frames of the camera that is rotating by Fig.4, 5 and 6. Here, Fig.4 indicates a tracking result obtained by our proposed method. Fig.5 does that tracking results by each element of our proposed method, namely, the tracking results obtained by the color data (upper) and those by the distance data (lower), respectively. Fig.6 does a tracking result obtained by the frame subtraction method. Red (shown in Fig.4 and 5 (lower)) and white (shown in Fig.5 (upper) and 6) squares in the sequences of the images in Fig.4, 5 and 6 means the detection result of the person’s head. The most right images in Fig.4, 5 and 6 are the detection results after the camera stops the pan motion, respectively.

From Fig.4, using both distance and color data of the targeted person is effective for detecting the targeted person while the camera is moving. The method is robust against change of the illumination condition, since the criteria of the method, that is the distance or the Hue value of the color, are robust against that [12],[13],[14]. Moreover, since the moving camera can distinguish the targeted person from the background by distance data, it can detect the target even if the colors of the background include a skin color.

We see that it is difficult to detect the person’s silhouette by the frame subtraction from the sequences of the detection result. The camera detects the person only while the camera is not moving. In the frame subtraction, when the camera is moving, brightness of each pixel changes by the camera motion as shown in Fig.6. Therefore, the subtraction result is noisy and not enough to detect the silhouette of the targeted person.

4. EXPERIMENT

This section describes the human tracking experiment using the proposed cooperative system. We state the experimental conditions such as the experimental environment, the layout of the fixed cameras and the moving camera, and the trajectory of the targeted person. Then, we show the experimental results to illustrate the feasibility of our method.

4.1 Experimental Conditions

We show the experimental environment and the trajectory of the targeted person in Fig.7, where this environment consists of two fixed cameras, one moving camera, and one server machine. The length and width of the experimental environment are 4.5[m] and 7[m]. We always use a right hand coordinate system in the experiment. We set the upper

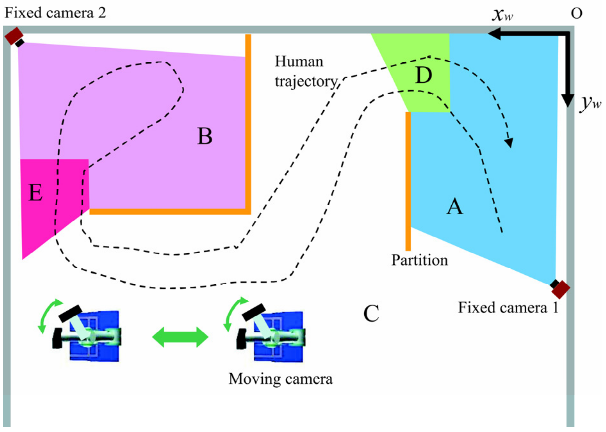

Fig.7. Experimental environment

right hand corner of the environment the origin of the world coordinate system, of which X-axis and Y-axis are x_w and y_w in Fig.7, respectively. Fixed camera 1 and 2 are located as shown in Fig.7. The height of the fixed cameras is set 1.8[m] from the floor. We use a mobile manipulator that the camera unit is attached to the end of the manipulator as the moving camera[7] in this experiment. We apply a trinocular camera unit "Digiclops Stereo Vision" as the fixed and the moving cameras.

Here, we first divide the whole environment into the following five areas.

Area A: Observable area of Fixed Camera 1
Area B: Observable area of Fixed Camera 2
Area C: Area unobservable of Fixed Camera 1 and 2 (Observable area only with Moving Camera)
Area D: Boundary area between observable and unobservable areas of Fixed Camera 1
Area E: Boundary area between observable and unobservable areas of Fixed Camera 2

In this experiment, the targeted person walks in the following sequence:

Area A $\rightarrow$ Area D $\rightarrow$ Area C $\rightarrow$ Area E $\rightarrow$ Area B $\rightarrow$ Area E $\rightarrow$ Area C $\rightarrow$ Area D $\rightarrow$ Area A

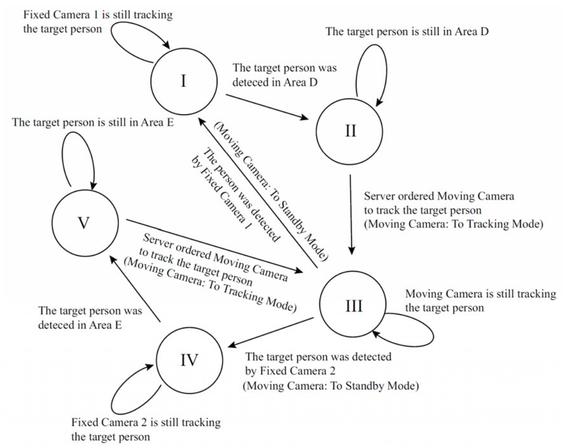

Fig.8. State transition diagram

Therefore the following five states can be considered as shown in Fig.8.

State I: Fixed Camera 1 is tracking the targeted person
State II: The targeted person is in Area D
State III: Moving Camera is tracking the targeted person
State IV: Fixed Camera 2 is tracking the targeted person
State V: The targeted person is in Area E

4.2 Experimental Results

We experimented on cooperative monitoring of human behavior in our experimental environment. The server machine records the trajectory of the targeted person measured by the cooperative system. The server can visualize its tracking result in real time. We show in Fig.9 the experimental result and in Fig.10 real-time visualization of monitoring human behavior by server.

We see that Fig.9 has the good result except some points of the moving camera parts. The reason why these points are outside the area of the moving camera results from the fact that its self-position was not identified accurately. The moving camera estimates itself position with several errors because of the slip on the floor, the counting error of its encoder, the self-calibration error in the environment, and so on. Therefore, this experiment was good as human tracking in each camera, however this system was not complete as a cooperative system since by the above fact the transformation matrix from the moving camera coordinate into the world coordinate was not correctly obtained. The other method for pose-calibration of the moving camera by itself or by the other fixed or moving cameras is required to improve the accuracy of the human tracking.

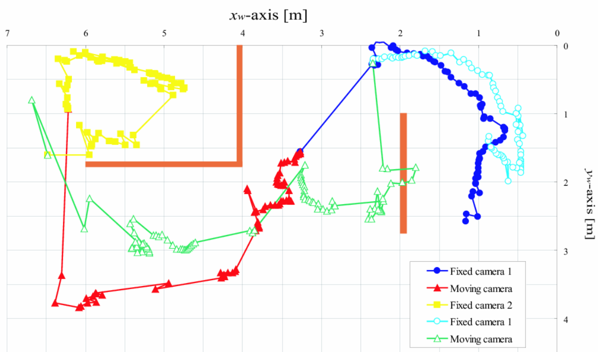

Fig.9. Experimental result

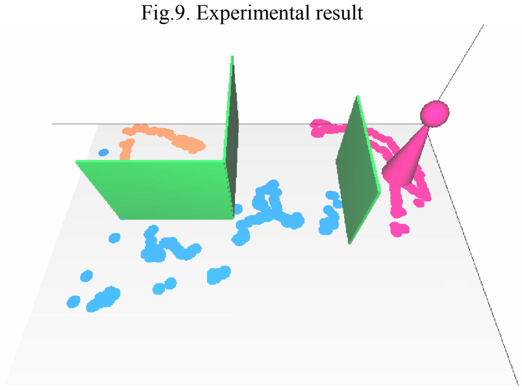
Fig.10. Visualized result of human tracking

5. CONCLUSION

In this paper, we have proposed a cooperative monitoring system by using fixed and moving cameras. First, we stated the cooperative system for monitoring. Then, we introduced a procedure of cooperation of the fixed and moving cameras. In particular, for monitoring a targeted person by the moving camera, we have applied the combination of the distance data from the camera to the targeted person and the color data of the person's head for detection of that targeted person. This detection method is robust against change of the illumination of the background and the motion of the camera. We finally carried out the person-tracking experiment using two fixed cameras and one moving camera.

We have obtained a comparatively good result in this experiment. Particularly, we have confirmed that human tracking is successful in an area unobservable of two fixed cameras by using a moving camera.

As future work, we are planning to correct the position data of a targeted person by correcting the self-position data of the moving camera sequentially or with some interval.

ACKNOWLEDGMENT

This work was supported by MEXT under Grand-in-Aid for Creative Scientific Research (Project No. 13GS0018).

REFERENCES

[1] J. H. Lee, K. Morioka, N. Ando and H. Hashimoto; Cooperation of Distributed Intelligent Sensors in Intelligent Environment, *IEEE/ASME Transactions on Mechatronics,* Vol. 9, No. 3, pp. 535 - 543 (2004)

[2] K. Morioka, J. H. Lee and H. Hashimoto; Human-Following Mobile Robot in a Distributed Intelligent Sensor Network, *IEEE Transactions on Industrial Electronics,* Vol. 51, No. 1, pp. 229 - 237 (2004)

[3] T. Matsuyama; Cooperative Distributed Vision: Dynamic Integration of Visual Perception, Action, and Communication, *Proc. of Image Understanding Workshop,* pp. 365–384 (1998)

[4] T. Matsuyama, S. Hiura, T. Wada, K. Murase and A. Toshioka; Dynamic Memory Architecture for Real Time Integration of Visual Perception, Camera Action, and Network Communication, *Proc. of IEEE Conference on Computer Vision and Pattern Recognition,* Vol. 2, pp.728-735 (2000)

[5] A. Nakazawa, H. Kato and S. Inokuchi; Human Tracking Using Distributed Vision Systems, *Proc. of 14th International Conference on Pattern Recognition,* Vol. 1, pp. 593–596 (1998)

[6] H. Kwon, Y. Yoon, J. B. Park and A. C. Kak; Person Tracking with a Mobile Robot using Two Uncalibrated Independently Moving Cameras, *Proc. of the 2005 IEEE International Conference on Robotics and Automation,* pp. 2888–2894 (2005)

[7] Y. Mae, N. Sasao, K. Inoue and T. Arai; Person Detection by Mobile-Manipulator for Monitoring, *Proc. of SICE Annual Conference 2003,* pp. 2845–2850 (2003)

[8] N. Matsuhira and H. Mizoguchi; Development of Life Support Robots that Coexist in Harmony with People, *Proc. of 36th International Symposium on Robotics,* CD-ROM (2005)

[9] E. Pagello, G. Clemente, A. D'Angelo, E. Menegatti, A. Pretto and S. Tonello; A Surveillance System based on a Distributed Vision System cooperating with a Mobile Robot, *Proc. of 36th International Symposium on Robotics,* CD-ROM (2005)

[10] Y. Wang and M. Hirakawa; Tracking Moving Objects in a Video System with Camera Motion, *IPSJ SIG Technical Report,* 2006 - CVIM - 152, pp. 129–136 (2006)

[11] W. Zajdel, Z. Zivkovic and B. Krose; Keeping Track of Humans: Have I Seen This Person Before ?, *Proc. of the 2005 IEEE International Conference on Robotics and Automation,* pp. 2093-2098 (2005).

[12] C. Eveland, K. Konolige and R. C. Bolles; Background Modeling for Segmentation of Video-Rate Stereo Sequences, *Proc. of IEEE Computer Society Conference on Computer Vision and Pattern Recognition,* pp. 266–271 (1998)

[13] J. Iio and K. Hiyane; Human Detection System using Robust Silhouette Extraction from Surveillance Images, *Proc. of 2nd International Symposium on Systems and Human Science,* CD-ROM (2005)

[14] S. L. Phung, A. Bouzerdoum and D. Chai; Skin Segmentation Using Color Pixel Classification: Analysis and Comparision, *IEEE Transaction on Pattern analysis and machine intelligence,* Vol. 27, No. 1, pp. 148– 154 (2005)