Indoor Positioning System Based on Distributed Camera Sensor Networks for Mobile Robot

Yonghoon Ji[1], Atsushi Yamashita[1], and Hajime Asama[1]

School of Engineering, The University of Tokyo, Japan,
t{ji, yamashita, asama}@robot.t.u-tokyo.ac.jp

Abstract. An importance of accurate position estimation in the field of mobile robot navigation cannot be overemphasized. In case of an outdoor environment, a global positioning system (GPS) is widely used to measure the position of moving objects. However, the satellite based GPS does not work indoors. In this paper, we propose a novel indoor positioning system (IPS) that uses calibrated camera sensors and 3D map information. The IPS information is obtained by generating a bird's-eye image from multiple camera images; thus, our proposed IPS can provide accurate position information when the moving object is detected from multiple camera views. We evaluate the proposed IPS in a real environment in a wireless camera sensor network. The results demonstrate that the proposed IPS based on the camera sensor network can provide accurate position information of moving objects.

Keywords: global positioning system, indoor positioning system, camera network, mobile robot

1 Introduction

This paper proposes an indoor positioning system (IPS) that uses calibrated camera sensor networks for mobile robot navigation. Most navigation functions that allow mobile robots to operate in indoor environments were performed based on only map information built by simultaneous localization and mapping (SLAM) schemes. Recently, however, robots are expected to be operated in human-robot coexistence environments. The map information is static; hence it is hard to deal with human-robot coexistence environments since it cannot reflect dynamic changes in the environment.

On the other hand, a distributed camera network system can monitor what is occurring in the environment and many automatic calibration schemes for such systems have been proposed in [8, 9]. Such system is able to manage environmental changes (e.g., moving objects) by processing image data from the distributed camera networks in real time; therefore, many of the problems encountered by classical mobile robot navigation can be improved by integrating both the static map information and the dynamic information from the calibrated camera sensor networks. The satellite based global positioning system (GPS) can greatly improve solutions to the positioning problem in an outdoor environment; however,

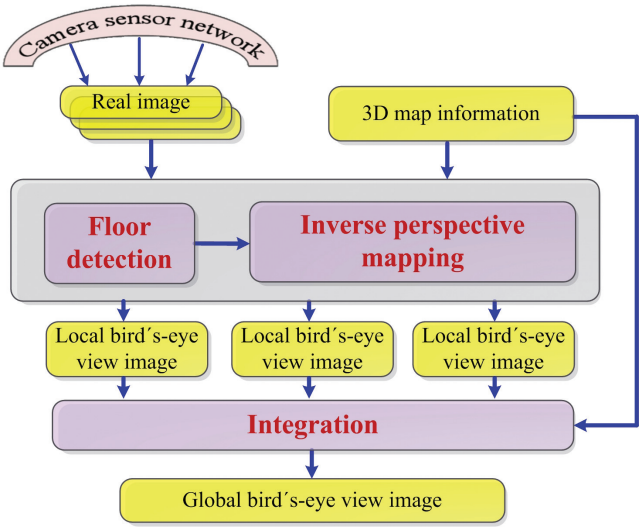

Fig. 1. Overall process to generate global bird's-eye image.

the GPS does not work indoors. In this respect, many IPS technologies that use different signals (e.g., Wi-Fi, electromagnetic, radio, and so on) are developed [4, 3]. However, many limitations such as the strength of the signal, accuracy, multi reflection phenomenon remain unanswered.

Therefore, in order to realize a reliable IPS, we propose a novel approach that uses both distributed camera sensor networks and 3D map information of the environment. The 3D map information contains the information for the entire environment based on the world coordinate frame, and thus it can provide information on any point. Such map information can be easily obtained from the blueprint of the artificial environment (e.g., CAD data) or a traditional SLAM schemes. In addition, generating a bird's-eye image from camera networks is one of the effective methods to provide IPS in the case of a typical indoor environment. A bird's-eye image is an elevated view of the environment. It is able to provide sufficient information when mobile robots navigate in a typical indoor environment because the ground is flat, and thus the height can be ignored. In our research, therefore, the position signal is calculated by using overlapping zones, which are detected from the generated bird's-eye image.

The remainder of this paper is organized as follows. Section 2 describes the generation method of the global bird's-eye image to generate IPS information based on the calibrated camera network system in detail. Then, positioning method for the IPS information with the generated global bird's-eye image is presented in section 3. The proposed IPS is validated with experiments for path planning of the mobile robot in sections 4. Finally, section 5 gives the conclusions.

2 Bird's-eye Image Generation

In this paper, to facilitate the monitoring of the entire environment, each of the calibrated camera images is transformed to local bird's-eye images that are taken from a user-define virtual camera, and then integrated into a global bird's-eye image. Figure 1 shows the overall process to generate the global bird's-eye image. Here, by way of example, three image data from camera networks are considered. First, floor detection process is performed to remove areas other than the floor using real camera images and predicted depth images generated from the camera poses (i.e., calibrated external camera parameters) and the 3D map information. Next, each of the generated floor images are converted into corresponding local bird's-eye images through an inverse perspective mapping and integrated into a global bird's-eye image. One of the notable features of this processing is that, with a distributed camera system, with accurately calibrated parameters based on the 3D map information, it is possible to generate RGB-D images (i.e., a RGB color image with a depth image) which provides 3D information, whereas only 2D image data cannot by itself provide 3D information. In this approach, the depth image generated from the 3D map information cannot mirror the dynamic information in real time because it is based on a static model; however, the RGB image is updated in real time. Therefore, even if this system cannot provide real time 3D information, it is able to handle 3D-like processing by utilizing overlapping areas from multiple camera images on an assumption that the ground is flat. Finally, the system performs positioning of IPS signal by using the generated global bird's-eye images.

2.1 Floor Detection

In order to detect the floor area in each of the real camera images captured by the camera sensor networks, first, height images $\boldsymbol{I}^{H}_{R(\boldsymbol{w})}$ are generated from both, the camera poses $\boldsymbol{w} = [x_c\ y_c\ z_c\ \psi_c\ \theta_c\ \phi_c]^\top$ and the 3D map information. Then, the real camera images $\boldsymbol{I}_{R(\boldsymbol{w})}$ (Fig. 2 (a)) are converted into floor images $\boldsymbol{I}^{F}_{R(\boldsymbol{w})}$ (Fig. 2 (d)) using the height images $\boldsymbol{I}^{H}_{R(\boldsymbol{w})}$ (Fig. 2 (c)) as bounding values:

$$\boldsymbol{I}^{F}_{R(\boldsymbol{w}^{(k)})}(u,v)=\begin{cases}\boldsymbol{I}_{R(\boldsymbol{w}^{(k)})}(u,v) & \boldsymbol{I}^{H}_{R(\boldsymbol{w}^{(k)})}(u,v)<0+\varepsilon \\ 0 & \text{otherwise}\end{cases}, \tag{1}$$

where k represents camera index. ε denotes the error constant which is added to take the camera calibration error into consideration. Thus, Eq. (1) removes regions other than the floor or those which heights are close to zero (the height of the floor surface is assumed to be zero in this study) from the real camera images $\boldsymbol{I}_{R(\boldsymbol{w})}$. It is reasonable to remove areas other than the floor given that the moving objects generally walk on the floor in a typical indoor environment.

The detailed process for the floor detection for the real camera image is as follows. First, each depth image $\boldsymbol{I}^{D}_{R(\boldsymbol{w})}$ is generated from the camera poses and the 3D map information as shown in Fig. 2 (b). Here, an important point to emphasize is that, with a distributed camera system with accurately calibrated

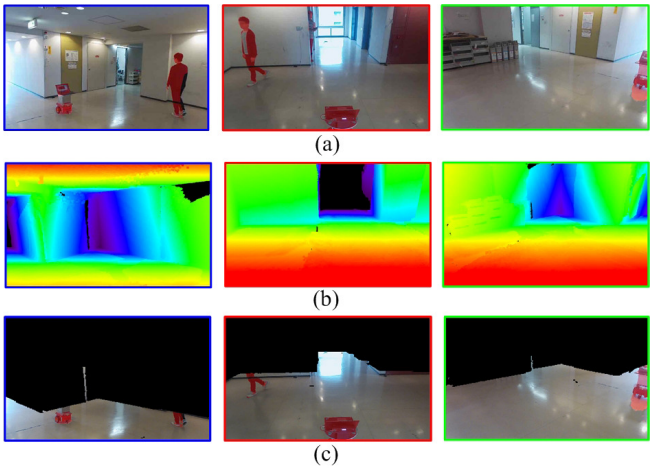

Fig. 2. Floor detection process: (a) real camera images $\boldsymbol{I}_{R(\boldsymbol{w})}$ captured from camera sensor networks, (b) predicted depth images $\boldsymbol{I}^{D}_{R(\boldsymbol{w})}$ that correspond to real camera images $\boldsymbol{I}_{R(\boldsymbol{w})}$ generated by 3D map information, (c) height images $\boldsymbol{I}^{H}_{R(\boldsymbol{w})}$ that correspond to real camera images $\boldsymbol{I}_{R(\boldsymbol{w})}$, and (d) floor images $\boldsymbol{I}^{F}_{R(\boldsymbol{w})}$ where regions other than the floor have been removed.

parameters and 3D map information, it is possible to generate RGB-D-like images even if a typical optical camera cannot directly obtain the depth information as mentioned before.

Next, each of height images $\boldsymbol{I}^{H}_{R(\boldsymbol{w})}$ (Fig. 2 (c)) that contain corresponding height information for each pixel are generated as follow:

$$\begin{aligned}\boldsymbol{I}^{H}_{R(\boldsymbol{w})}(u,v) = z_c - {}^{C}x\sin\theta_c + {}^{C}y\cos\theta_c\sin\psi_c \\ + {}^{C}z\cos\theta_c\sin\psi_c, \qquad (2)\end{aligned}$$

$$ {}^{C}x = \boldsymbol{I}^{D}_{R(\boldsymbol{w})}(u,v), \qquad (3)$$

$$ {}^{C}y = \frac{(u-c_u)\boldsymbol{I}^{D}_{R(\boldsymbol{w})}(u,v)}{f_u}, \qquad (4)$$

$$ {}^{C}z = \frac{(c_v-v)\boldsymbol{I}^{D}_{R(\boldsymbol{w})}(u,v)}{f_v}. \qquad (5)$$

Equations (3)–(5) convert 2D pixel information (u, v) to 3D data $({}^{C}x, {}^{C}y, {}^{C}z)$ based on the local camera coordinate frame using the pixel information of the depth image $\boldsymbol{I}^{D}_{R(\boldsymbol{w})}$. Here, the superscript C indicates the local camera coordinate frame. $\boldsymbol{w} = [x_c\ y_c\ z_c\ \psi_c\ \theta_c\ \phi_c]^{\top}$ represents the camera pose with respect to the world coordinate frame $\{\boldsymbol{W}\}$. f_u and f_v denote focal lengths. Equation (2) refers to the coordinate transformation from camera coordinate frame to world coordinate frame $\{\boldsymbol{W}\}$ for the height value. Finally, the floor images $\boldsymbol{I}^{F}_{R(\boldsymbol{w})}$ (Fig. 2 (d)) are generated by applying Eq. (1). The coordinate system adopted

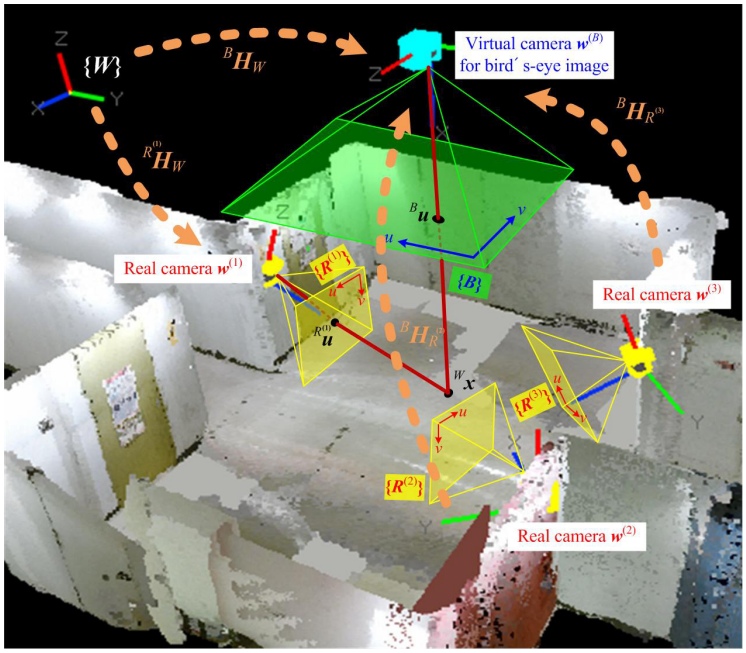

Fig. 3. Perspective transformations between floor plane, real camera images and bird's-eye image for virtual camera.

in this study is described in detail in Fig. 3. Note that the optical axis is defined as x-axis in the camera coordinate frame in this paper.

2.2 Inverse Perspective Mapping

In order to generate local bird's-eye images from each camera, an inverse perspective mapping is exploited. The inverse perspective mapping removes perspective effects under the assumption that the ground is flat; therefore making environment monitoring more efficient [1]. In this study, the inverse perspective mapping is applied to each of the generated floor images $\boldsymbol{I}^F_{R(\boldsymbol{w})}$. This is done as it is assumed that the mobile robot navigates on the floor area in a typical indoor environment.

The bird's-eye image is generated by image processing based on the principle of geometrical perspective transformation for the captured real image. The principle used to convert real camera images to bird's-eye images for a virtual viewpoint is as follows. Figure 3 illustrates perspective transformations between each real camera and a user-define virtual camera that represents the relationship between a world coordinate frame $\{\boldsymbol{W}\}$, real camera images coordinate frames $\{\boldsymbol{R}\}$, and virtual bird's-eye image coordinate frame $\{\boldsymbol{B}\}$. Each relationship is described using transformation matrices. First, a relational expression between

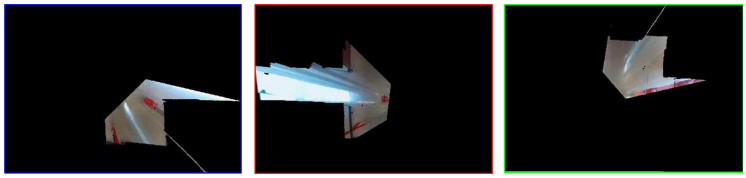

Fig. 4. Local bird's-eye images $\boldsymbol{I}^F_{B(\boldsymbol{w})}$ corresponding to floor images $\boldsymbol{I}^F_{R(\boldsymbol{w})}$.

${}^W\boldsymbol{x} = [x\ y\ z\ 1]^\top$ and ${}^R\boldsymbol{u} = [{}^Ru\ {}^Rv\ s]^\top$ is defined as follow:

$$
\begin{aligned}
{}^R\boldsymbol{u} &= {}^R\boldsymbol{Q}_W\,{}^W\boldsymbol{x} \\
\begin{bmatrix} {}^Ru \\ {}^Rv \\ s \end{bmatrix} &= \begin{bmatrix} h_{11}\ h_{12}\ h_{13}\ h_{14} \\ h_{21}\ h_{22}\ h_{23}\ h_{24} \\ h_{31}\ h_{32}\ h_{33}\ h_{34} \end{bmatrix} \begin{bmatrix} {}^Wx \\ {}^Wy \\ {}^Wz \\ 1 \end{bmatrix},
\end{aligned} \tag{6}
$$

where s denotes a scale factor. The expression can be further simplified under the assumption that the height of the floor surface is zero (i.e., ${}^W\boldsymbol{z} = 0$), then:

$$
\begin{aligned}
{}^R\boldsymbol{u} &= {}^R\boldsymbol{H}_W\,{}^W\tilde{\boldsymbol{x}} \\
\begin{bmatrix} {}^Ru \\ {}^Rv \\ s \end{bmatrix} &= \begin{bmatrix} h_{11}\ h_{12}\ h_{14} \\ h_{21}\ h_{22}\ h_{24} \\ h_{31}\ h_{32}\ h_{34} \end{bmatrix} \begin{bmatrix} {}^Wx \\ {}^Wy \\ 1 \end{bmatrix},
\end{aligned} \tag{7}
$$

where ${}^W\tilde{\boldsymbol{x}} = [{}^Wx\ {}^Wy\ 1]^\top$ is the position vector on the world coordinate frame excepting the z element. The matrix ${}^R\boldsymbol{H}_W$ is called homography matrix in the geometric field. It is possible to project one surface onto another through this homography matrix.

In the same manner, a relational expression between ${}^W\tilde{\boldsymbol{x}} = [{}^Wx\ {}^Wy\ 1]^\top$ and ${}^B\boldsymbol{u} = [{}^Bu\ {}^Bv\ s]^\top$ is also defined:

$$
{}^B\boldsymbol{u} = {}^B\boldsymbol{H}_W\,{}^W\tilde{\boldsymbol{x}}, \tag{8}
$$

where ${}^B\boldsymbol{H}_W$ is the homography matrix between the world coordinate frame $\{\boldsymbol{W}\}$ and the bird's-eye image coordinate frame $\{\boldsymbol{B}\}$ when the height of the projection plane is zero.

Finally, substituting Eq. (8) into Eq. (7) yields the relation between the real image coordinate frame $\{\boldsymbol{R}\}$ and the bird's-eye image coordinate frame $\{\boldsymbol{B}\}$:

$$
{}^B\boldsymbol{u} = {}^B\boldsymbol{H}_R\,{}^R\boldsymbol{u}. \tag{9}
$$

As shown in Eq. (9), the homography matrix ${}^B\boldsymbol{H}_R$ performs inverse perspective mapping (i.e., the transforming the real camera image into the local bird's-eye image). It is very easy to calibrate the 3×3 homography matrix ${}^B\boldsymbol{H}_R$ by randomly selecting more than 4 points (i.e., at least four ${}^W\boldsymbol{x}$) that are commonly observed in the real camera image $\{\boldsymbol{R}\}$ and the bird's-eye image $\{\boldsymbol{B}\}$

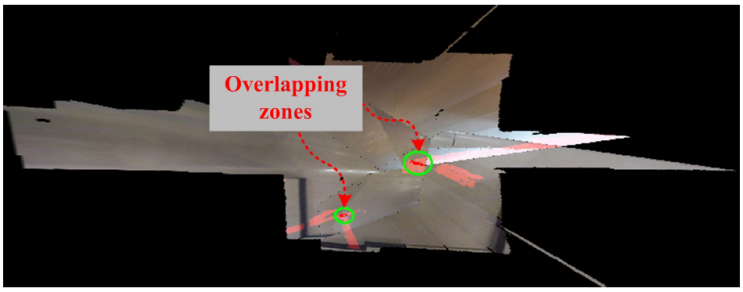

Fig. 5. Integrated global bird's-eye images ${}^{G}\boldsymbol{I}^{F}_{B(\boldsymbol{w})}$.

because pre-given 3D map information contains the information for the entire environment based on the world coordinate frame $\{\boldsymbol{W}\}$, and thus it can provide information on any point.

Figure 4 shows each local bird's-eye image $\boldsymbol{I}^{F}_{B(\boldsymbol{w})}$ that is generated by applying each calibrated homography matrix ${}^{B}\boldsymbol{H}_{R}$ to the floor image $\boldsymbol{I}^{F}_{R(\boldsymbol{w})}$. In this case, the virtual camera pose was set to the high position at the center of the entire environment.

2.3 Integration of Local Bird's-eye Images

After generating each of the local bird's-eye images $\boldsymbol{I}^{F}_{B(\boldsymbol{w})}$ by applying the inverse perspective mapping, these images are combined together to create a global bird's-eye image ${}^{G}\boldsymbol{I}^{F}_{B(\boldsymbol{w})}$ that is viewed from direct observation at high location, as follow:

$$ {}^{G}\boldsymbol{I}^{F}_{B(\boldsymbol{w})}(u,v) = {}^{G}\boldsymbol{I}^{F}_{B(\boldsymbol{w}^{(k^*)})}({}^{B}u^*, {}^{B}v^*), \tag{10} $$

$$ \begin{bmatrix} {}^{B}\boldsymbol{u}^* \\ {}^{B}\boldsymbol{v}^* \\ 1 \end{bmatrix} = {}^{B}\boldsymbol{H}_{R^{(k^*)}} \begin{bmatrix} {}^{R}\boldsymbol{u}^* \\ {}^{R}\boldsymbol{v}^* \\ 1 \end{bmatrix}, \tag{11} $$

$$ (k^*, {}^{R}\boldsymbol{u}^*, {}^{R}\boldsymbol{v}^*) = \underset{k,u,v}{\arg\min} \left[\boldsymbol{I}^{D}_{R(\boldsymbol{w}^{(k)})}(u,v) \right]. \tag{12} $$

Equations (10)–(12) mean that when the same point is observed by multiple cameras, the information encoded from the pixel which has the shortest distance to its camera's optical axis is used. Here, k refers to the camera index (i.e., 1–3 in this example). To calculate the shortest distance, the depth images $\boldsymbol{I}^{D}_{R(\boldsymbol{w})}$ for each camera are utilized again here. This is justified as the distortion of the bird's-eye image is smaller when the physical distance in real space is closer.

Figure 5 shows the combined global bird's-eye image ${}^{G}\boldsymbol{I}^{F}_{B(\boldsymbol{w})}$ that is generated by applying Eqs. (10)–(12). Using this approach, it is possible to observe the entire area of the floor that is captured by the camera sensor networks intuitively. Furthermore, the proposed global bird's-eye image can be of great service to moving object's positioning. This will be presented in the following section.

3 Positioning method for IPS

This section describes positioning method for the proposed IPS by using the generated global bird's-eye image. In this paper, the background subtraction method is used for recognizing moving objects in the image planes. Among the many background subtraction methods, mixture Gaussian model-based method has been used for this study [12]. Red areas in Figs. 2 (a), 2 (d), 4, and 5 show each detection result after applying the background subtraction. As we can see in Fig. 4, the detected objects from single cameras are significantly distorted in the local bird's-eye images because 3D viewpoint transformation is forcibly preformed for the 2D image data in order to generate the local bird's-eye images. These distortion effects are obviously inefficient in terms of reliable obstacle detection.

In Fig. 5, the dark red areas are the overlapping zones that represent detected moving objects from the multiple camera network. In this study, the distortion effect is solved by taking advantage of these overlapping areas. Thus, if the moving objects are observed from two or more cameras, the distortion effects can be corrected as shown in Fig. 6. 3D-like positioning can be performed in this case without a tedious and convoluted stereo measurement task when the height of the floor can be assumed constant value through the 3D map information.

The principle used to correct these distortions is illustrated in Fig. 6. In contrast to stereo cameras, a single camera cannot measure the distance to the object directly. However, if the object is observed from (i.e., detected from background subtraction in this study) multiple cameras from different viewpoints, then the uncertainty of the distance (i.e., existence region) will be significantly reduced. Note that there is a thread of connections between this principle and well-known monocular SLAM with a single camera [5]. In the global bird's-eye image with the proposed combined method can remove distortion effects very intuitively, whereas monocular SLAM schemes use convoluted probabilistic processes.

The detected results are composed of pixel coordinates in the global bird's-eye image frame $\{\boldsymbol{B}\}$; thus, the detected regions ${}^{B}\boldsymbol{u}_{\mathrm{IPS}}$ as pixel coordinates should be transformed to those of real coordinates ${}^{W}\tilde{\boldsymbol{x}}_{\mathrm{IPS}}$ through the homography

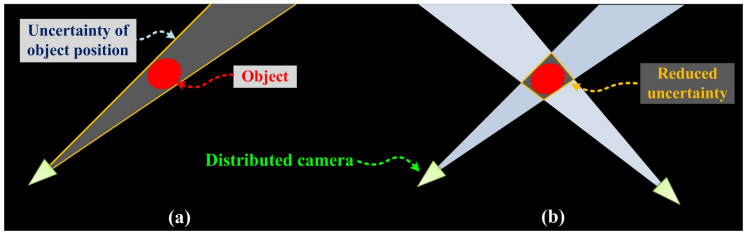

Fig. 6. Principle of distortion removal based on multiple camera observations from different viewpoints: (a) object detected from one camera, (b) object detected from two cameras.

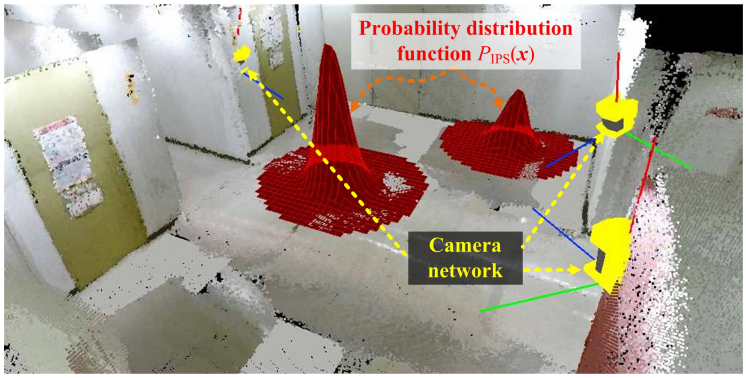

Fig. 7. Probability distribution $P_{\mathrm{IPS}}(\boldsymbol{x})$ for the existence of moving object detected by proposed IPS system.

matrix ${}^{W}\boldsymbol{H}_{B}$:

$$\begin{aligned} {}^{W}\tilde{\boldsymbol{x}}_{\mathrm{IPS}} &= \left({}^{B}\boldsymbol{H}_{W}\right)^{-1}{}^{B}\boldsymbol{u}_{\mathrm{IPS}} \\ &= {}^{W}\boldsymbol{H}_{B}{}^{B}\boldsymbol{u}_{\mathrm{IPS}}. \end{aligned} \tag{13}$$

Here, ${}^{B}\boldsymbol{H}_{W}$ is defined on user-defined internal parameter matrix and external parameter matrix which represents virtual camera pose for generating the bird's-eye image. Here, its 3$^{\mathrm{rd}}$ row vector is excluded because the height of the ground plane is defined as zero. Next, the coordinates ${}^{W}\tilde{\boldsymbol{x}}_{\mathrm{IPS}}$ detected as the moving objects are converted into a probability distribution function through Kernel density estimation:

$$P_{\mathrm{IPS}}(\boldsymbol{x}) = \frac{1}{Nh^{D}}\sum_{i=1}^{N}\boldsymbol{K}\left(\frac{\boldsymbol{x}-\boldsymbol{x}_{\mathrm{IPS}}^{(i)}}{h}\right), \tag{14}$$

$$\boldsymbol{K}(\boldsymbol{x}) = \frac{1}{(2\pi)^{D/2}}\exp\left(-\frac{1}{2}\boldsymbol{x}^{\top}\boldsymbol{x}\right), \tag{15}$$

where $\boldsymbol{x}_{\mathrm{IPS}}^{(i)}$ and N denote the coordinates of the detected zone and the number of detected pixels ${}^{B}\boldsymbol{u}_{\mathrm{IPS}}$ in the global bird's-eye image. From here, the superscript W which represents world coordinate frame is omitted. h and D are a smoothing parameter and the point dimensions (two in this case), respectively. Equation (15) represents typical $\boldsymbol{O}$ mean and $\boldsymbol{I}$ covariance Gaussian Kernel function $\boldsymbol{K}(\cdot)$ which is applied to manage several noises owing to errors from camera calibration, background subtraction, and so on. Figure 7 shows generated final IPS information $P_{\mathrm{IPS}}(\boldsymbol{x})$, which represents the existence probability distributions of the moving objects detected by multiple camera sensor networks.

4 Experimental result for safe path generation

This section presents an experimental result for safe path generation in a human-robot coexistence environment based on a gradient method with 3D map information and the information from camera sensor networks. Most obstacle avoidance schemes were based on processing on-board sensor (i.e., a sensor mounted on robot's body) information so far, and thus these cannot deal with invisible obstacles because of the sensing scope. In this study, therefore, in order to take invisible dynamic obstacles into consideration, occluded obstacles are detected by processing the global bird's-eye view image generated from camera sensor networks installed in the human-robot coexistence environment. In other words, this section assumes that generated IPS information represents the positions of the obstacles. The possibility of collision is basically reflected to path generation and motion control algorithms. Major scope in this section is generating safe paths for the mobile robot, taking not only visible obstacles but also occluded moving obstacles into consideration. Most path planning methods generate optimal paths based on only visible obstacles and the optimality is basically defined in terms of distance from a current robot position to a goal position. However, the shortest path might be not the safest path for the robot motion. The path should reflect the risk of possible collision with occluded invisible obstacles. To this end, the proposed path generation scheme calculates posterior existence probabilities of dynamic obstacles including occluded ones to redefine the cost function of a gradient method [13]. By exploiting the redefined cost function, safe path taking dynamic obstacles' existence probabilities into account is generated.

The gradient method generates a minimum distance path without local minima problem, and it is the most widely used method. The original gradient method generates an optimal path based on intrinsic costs and adjacency costs which are allocated to every grids of the map information. The intrinsic costs are assigned for the distance from the static obstacles which are represented in the map information and the adjacency costs are assigned for the distance from the goal position. In addition to these costs, the modified gradient method proposed in this study calculates an additional risk costs that correspond to moving obstacles which are detected by the global bird's-eye view image generated from the multiple camera images. This is done in order to perform safer path planning also considering occluded zones in real time. Here, the probability distribution $P_{\mathrm{IPS}}(\boldsymbol{x})$ (i.e., IPS information) which is computed by Eqs. (14) and (15) can be directly exploited, as the additional risk cost relates to the existence probability for the moving obstacles in the entire environment. What this entail is that positions which have high value of $P_{\mathrm{IPS}}(\boldsymbol{x})$ are more likely have a moving obstacle.

The experiments were conducted under the same conditions using two different methods: the conventional gradient method considering only the intrinsic cost and the adjacency cost and the proposed modified gradient method considering the additional risk cost representing the moving obstacles. The generated path using each method is illustrated in Fig. 8. The generated path using the conventional gradient method generates an unsafe path (blue lines in Fig. 8)

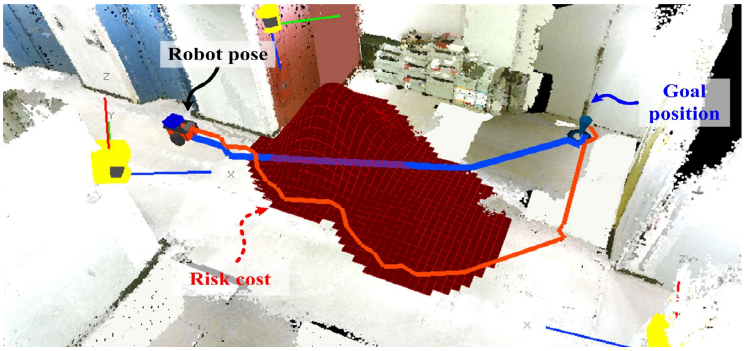

Fig. 8. Experimental results for path generation. Blue and red lines represent generated paths using typical gradient method and modified gradient method, respectively.

around the moving obstacles because the static model (i.e., the map information) cannot manage the dynamic information in real time, and thus it cannot be reflected in the intrinsic cost for the gradient method. On the other hand, the modified gradient method produced a safer path (red lines in Fig. 8) taking moving obstacles into account as the risk costs for the moving objects detected from the global bird's-eye view image were calculated and applied to the total navigation costs for optimal path generation. In conclusion, the proposed path planning scheme based on the modified gradient method considering occluded moving obstacles is expected to reduce collision risk in terms of the mobile robot's motion control.

5 Conclusion

This paper proposed a novel IPS that uses distributed camera sensor networks for the mobile robot navigation. In order to generate reliable IPS information from the camera networks, a novel method to generate a global bird's-eye image was proposed by using 3D map information. Here, homography matrices for transforming each real image to each local bird's-eye image were automatically calibrated using 3D map information which had all 3D coordinate data for the entire environment. The global bird's-eye view image generated from the multiple camera images was used to detect moving obstacles (i.e., generation of IPS information), and the following conclusion was drawn.

The typical path planning methods so far cannot manage occluded obstacles due to limitations of the on-board sensor's sensing scope. By using the information of the moving obstacles detected as an additional cost function to the gradient method, a safer path can be generated. Thus, mobile robots are expected to reduce their collision risks.

Future work will involve solving a problem of the optimal distributed camera placement by maximizing the coverage of the demands.

Acknowledgements

This work was in part supported by Tough Robotics Challenge, ImPACT Program (Impulsing Paradigm Change through Disruptive Technologies Program).

References

1. Bertozz, M., Broggi, A., and Fascioli, A.: Stereo Inverse Perspective Mapping: Theory and Applications. Image and Vision Computing, 16(8), 585–590 (1998)
2. Brščić, D. and Hashimoto, H.: Model Based Robot Localization Using Onboard and Distributed Laser Range Finders. Proceedings of the 2008 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS2008), 1154–1159 (2008)
3. Chang, N., Rashidzadeh, R., and Ahmadi, M.: Robust Indoor Positioning Using Differential Wi-Fi Access Points. IEEE Transactions on Consumer Electronics, 56(3), 1860–1867 (2010)
4. Curran, K., Furey, E., Lunney, T., Santos, J., Woods, D., and McCaughey, A.: An Evaluation of Indoor Location Determination Technologies. Journal of Location Based Services, 5(2), 61–78 (2011)
5. Davison, A.J., Reid, I.D., Molton, N.D., and Stasse, O.: MonoSLAM: Real-time Single Camera SLAM. IEEE Transactions on Pattern Analysis and Machine Intelligence, 29(6), 1052–1067 (2007)
6. Dellaert, F., Fox, D., Burgard, W., and Thrun, S.: Monte Carlo Localization for Mobile Robots. Proceedings of the 1999 IEEE International Conference on Robotics and Automation (ICRA1999), 1322–1328 (1999)
7. Jayasekara, P.G., Hashimoto, H., and Kubota, T.: Simultaneous Localization Assistance for Mobile Robot Navigation in Real, Populated Environments. SICE Journal of Control, Measurement, and System Integration, 5(6), 349–358 (2012)
8. Ji, Y., Yamashita, A., and Asama, H.: Automatic Camera Pose estimation Based on Textured 3D Map Information. Proceedings of the 2015 6th JSME/RMD International Conference on Advanced Mechatronics (ICAM2015), pp. 100–101 (2015)
9. Ji, Y., Yamashita, A., and Asama, H.: Automatic Calibration and Trajectory Reconstruction of Mobile Robot in Camera Sensor Network. Proceedings of the 11th Annual IEEE International Conference on Automation Science and Engineering (CASE2015), 206–211 (2015)
10. Rahimi, A., Dunagan, B., and Darrell, T.: Simultaneous Calibration and Tracking with a Network of Non-overlapping Sensors. Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR2004), I–287 (2004)
11. Reza, A.W. and Geok, T.K.: Investigation of Indoor Location Sensing Via RFID Reader Network Utilizing Grid Covering Algorithm. Wireless Personal Communications, 49(1), 67–80 (2009)
12. Stauffer, C. and Grimson, W.E.L.: Adaptive Background Mixture Models for Real-time Tracking. Proceedings of the 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR1999), 246–252 (1999)
13. Konolige, K.: A Gradient Method for Realtime Robot Control. Proceedings of the 2000 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS2000), 639–646 (2000)