Title	Superresolution modeling using an omnidirectional image sensor
Author(s)	Nagahara, Hajime; Yagi, Yasushi; Yachida, Masahiko
Citation	IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics. 2003, 33(4), p. 607-615
Version Type	VoR
URL	https://hdl.handle.net/11094/3124
rights	©2003 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.
Note	

Superresolution Modeling Using an Omnidirectional Image Sensor

Hajime Nagahara, Yasushi Yagi, and Masahiko Yachida

Abstract—Recently, many virtual reality and robotics applications have been called on to create virtual environments from real scenes. A catadioptric omnidirectional image sensor composed of a convex mirror can simultaneously observe a 360-degree field of view making it useful for modeling man-made environments such as rooms, corridors, and buildings, because any landmarks around the sensor can be taken in and tracked in its large field of view. However, the angular resolution of the omnidirectional image is low because of the large field of view captured. Hence, the resolution of surface texture patterns on the three–dimensional (3-D) scene model generated is not sufficient for monitoring details. To overcome this, we propose a high- resolution scene texture generation method that combines an omnidirectional image sequence using image mosaic and superresolution techniques.

Index Terms—Image mosaic, modeling, omnidirectional image sensor, SR modeling, superresolution.

I. INTRODUCTION

RECENTLY, many virtual reality and robotics applications have been called on to create virtual environments from real scenes. An omnidirectional image sensor can observe a $360°$ field of view making it useful for scene modeling. Landmarks around the sensor can be taken in and tracked in its large field of view.

Various omnidirectional image sensors have been proposed (see [1] for survey). In particular, catadioptric omnidirectional image sensors using convex mirrors [2], [3] are useful for modeling. These sensors can simultaneously capture omnidirectional information and can continuously observe objects while the camera moves around an environment. Another advantage is that the catadioptric omnidirectional image sensors are usually portable, and therefore an omnidirectional movie can be easily recorded while walking down a road, driving a car, taking a train, etc.

Many researchers have investigated several three–dimensional (3-D) modeling methods with omnidirectional image sensors [4]–[15]. In practice, however, the environmental model generated cannot be applied to the applications mentioned above because of the sensor's low angular resolution, and the image resolution of the observed surface texture is too low for monitoring details. In essence, catadioptric omnidirectional image sensors have the advantages of simultaneous omnidirectional sensing and easy handling, but the disadvantage of low angular resolution.

Besides using a catadioptric camera, another approach for obtaining high-resolution omnidirectional images is to rotate the camera around its vertical axis. Recently, Nayer has proposed a method to rotate the omnidirectional image sensor around an axis perpendicular to the lens axis [16]. We have also proposed a high-resolution omnidirectional imaging system that combines consecutive omnidirectional images obtained by rotating an omnidirectional image sensor around a lens axis [17], [18]. However, with these methods, the camera must remain motionless while recording images. Hence, these methods cannot be applied to an omnidirectional image sequence while the sensor moves along a path.

A third method that does not involve rotating the camera is the use of multiple cameras. Kawanishi *et al.* constructed the high-resolution omnidirectional image sensor [19] using hexagonal pyramidal mirrors and six cameras. The system can acquire a high resolution at a video-rate, however, they pointed out that alignment and calibration among each of the six cameras is difficult and not yet resolved. Further, it is difficult to build it into a compact system because the use of six cameras requires the same number of AD converters or recorders. Such a complex system is thus difficult to handle easily.

In this paper, we propose a resolution improvement method for a 3–D model generated by a catadioptric omnidirectional image sensor. The proposed method only requires an omnidirectional video stream with smooth sensor motion. The method improves the resolution of textures mapped on the geometrical surface model, using image mosaic and superresolution techniques, generated from an image sequence. Our modeling method can be applied not only to an omnidirectional image sensor but also to a standard camera. We call this concept superresolution modeling.

Many superresolution methods combining multiple low-resolution images have been proposed for improving resolution [20]–[22]. Schultz and Stevenson [23] have embedded a method in a Bayesian framework. Cheeseman [24] has proposed a method that utilizes parallel processing. Mann and Picard [25] have defined a transformation matrix between two arbitrary image coordinates that allows an image to be projected to another image plane. The superresolution is then done on one of the image planes. Their projective transformation was restricted to two cases: 1) camera rotated around a vertical axis passing through the center of the camera, and 2) The target object in the scene was just a plane. The superresolution

The authors are with the Graduate School of Engineering Science, Osaka University, Toyonaka, Osaka 560-8531, Japan; (e-mail: nagahara@yachi-lab.sys.es.osaka-u.ac.jp; y-yagi@sys.es.osaka-u.ac.jp; yachida@sys.es.osaka-u.ac.jp).

techniques discussed above, which we call superresolution imaging (as opposed to superresolution modeling), usually generate a high-resolution image on the projective image plane at a viewpoint where low-resolution input images have been captured. In contrast, our proposed method is superresolution on the 3-D real object surface. Our proposed method of "superresolution modeling" directly generates a high-resolution texture that is mapped on the model. The differences between these two concepts are described in detail in Section II.

Capel and Zisserman [26] estimated the resolution enhancement ratio (resulting image size) from the condition number of a matrix of input images and the number of input images captured by a common video camera. This is suitable for deciding the number of input images. However, it cannot be directly applied to our method, because the optical specifications of the catadioptric omnidirectional image sensors are special. For example, its input image resolution is not constant and is greatly changed by depression angles. The resolution on the planar model surface is also changed by the distance and posture of the model surface relative to the sensor.

In this paper, we define a weighting table that represents the resolution difference between the model surface and the input image captured by a catadioptric image sensor. The weighting table is used to precisely align input image textures to each model surface, and then used as the criterion for selecting input image textures for superresolution.

Fig. 1. Resolution density of virtual image with SR imaging and SR modeling.

II. SUPERRESOLUTION IMAGING VERSUS SUPERRESOLUTION MODELING

As shown in Fig. 1(a), superresolution imaging is the generation of a high-resolution image by projecting inputs to a target image plane. Usually, the target image plane is fixed at one point and is selected from viewpoints View1 where input images have been captured. Here, if we set the virtual viewpoint View2 across the real object surface as shown in Fig. 1(a), that part of the image resolution at the viewpoint View2 is decreasing and the distribution of the resolution in the high-resolution image is enlarging. This means that image resolution is localized inhomogeneously.

On the other hand, superresolution modeling is the generation of a surface model with a high-resolution texture. Input images are projected to a real object plane in a geometrical model and superresolution is done on this object plane. Therefore, we can optimize the texture resolution on the object plane in the geometrical model as shown in Fig. 1(b). Texture resolution is both high and also uniformly distributed.

Here, we analyzed which combination was better, superresolution imaging with geometrical modeling or superresolution modeling with geometrical modeling.

superresolution imaging with geometrical modeling is that the camera stops and captures low-resolution images and generates the high-resolution images at each point. Then, the 3-D geometrical model is generated from the high-resolution images at each point. For instance, to generate the correct superresolution image, we have to capture input images by rotating the camera. However, it loses the advantage that the catadioptric omnidirectional image sensor can simultaneously capture omnidirectional
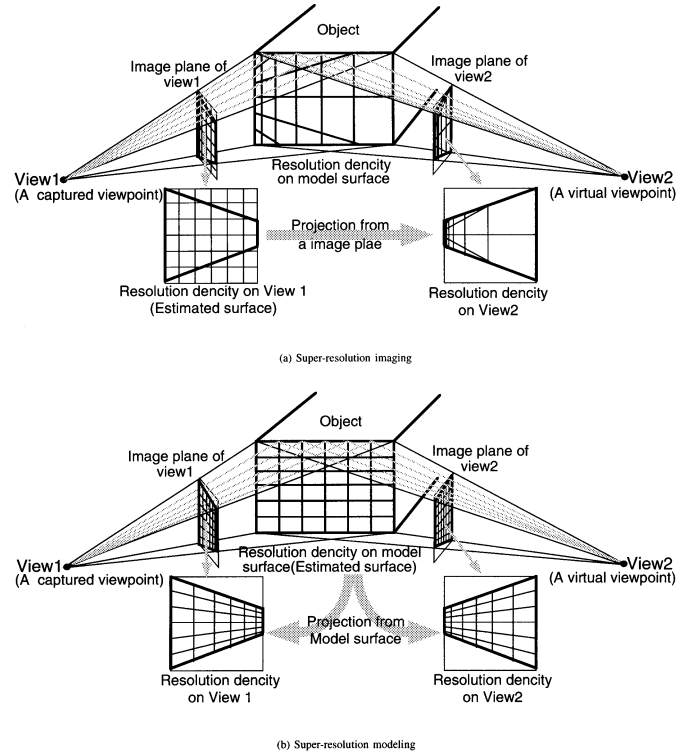
information and continuously observe objects. If input images are captured while the robot is moving, the position of features for geometrical modeling is not assured in the superresolution image, because absolute position alignments among input images are not required for superresolution, which requires only relative position alignment.

On the other hand, superresolution modeling with geometrical modeling first estimates a geometrical model and then estimates a high-resolution texture on a surface model. It can enhance a texture resolution from images with consecutive sensor motion and restrict the searching region on the matching process, because surface positions and sensor egomotion are already known by the geometrical modeling before estimating the high-resolution texture. We can estimate the precise egomotion and geometrical model using consecutive measurements of features in omnidirectional images while the camera is moving, by the least squares method.

Thus, we considered that superresolution modeling is suitable for combining with geometrical modeling. In this paper, we develop the superresolution modeling concept that estimates a superresolution texture mapped on the model surface. Our proposed method first makes a 3-D geometrical model from images captured with arbitrary smooth sensor motion, and enhances the surface texture by using the superresolution and image mosaic technique.

III. ASSUMPTION

The following properties of the environment are assumed for image analysis. The robot (sensor) moves in a man-made stationary environment such as a room or a corridor in which the
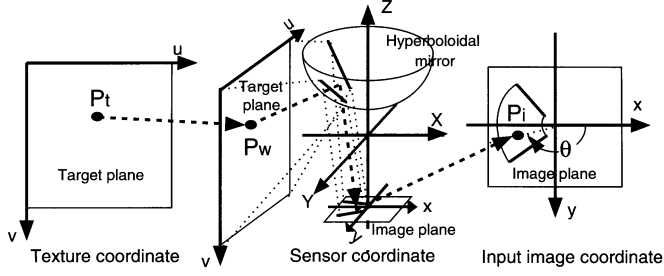
Fig. 2.   Hyperboloidal projection.

Fig. 3.   Alignment of texture images.

floor is horizontal and almost flat. The modeling targets are walls and static objects such as desks or shelves that have vertical planes. Lighting conditions are indirect.

IV. THREE-DIMENSIONAL GEOMETRICAL MODELING USING OMNIDIRECTIONAL IMAGE SENSOR

Our method of resolution improving is for a 3-D geometrical model generated from consecutive omnidirectional images captured by the HyperOmni Vision omnidirectional image sensor [2]. Fig. 2 shows the projective relationship for HyperOmni Vision. HyperOmni Vision, using a TV camera with its optical axis aligned with the axis of the hyperboloidal mirror, is a practical omnidirectional image sensor. With this structure, we can acquire a 360-degree view around the optical axis of the camera. A hyperboloidal mirror yields the image of a point in space on a vertical plane through the point P_w and its axis. This means that the angle in the image, which can be easily calculated as y/x, shows the azimuth angle θ of the point P_w in space (See Fig. 2). Also, it is easily understood that all points with the same azimuth in space appear on a radial line through the image center as shown in Fig. 2. Therefore, in the image, the vertical edges in the environment appear radially, and azimuth angles don't vary with changes in distance and height. Estimations of the position of vertical edges and the egomotion of the omnidirectional image sensor can be achieved by monitoring the locus of the azimuth angles of vertical edges while the omnidirectional image sensor is moving. By using this relation, various researchers have proposed several map generation methods [4]–[7]. Even with the unknown motion of the mobile robot, our proposed method can estimate the robot locations and generate the environmental map in real time [7].

Here, the point $P_w(X, Y, Z)$ on the sensor coordinate is projected onto the image point P_i at (x, y) as (1), (2), and (3). a, b, c are coefficients of the hyperboloidal mirror. They have relationships as $(X^2 + Y^2)/a^2 - Z^2/b^2 = -1$ and $c = (a^2 + b^2)^{1/2}$. Heights of vertical edges are easily calculated from (2) and (3) if the vertical edge positions (X, Y) are given. By using such a method, the 3-D geometrical surface model with all its surface texture can be generated. More details of the 3-D geometrical model generation are in our previous report [7].

V. RESOLUTION IMPROVING METHOD FOR OMNIDIRECTIONAL VISION BASED MODELING

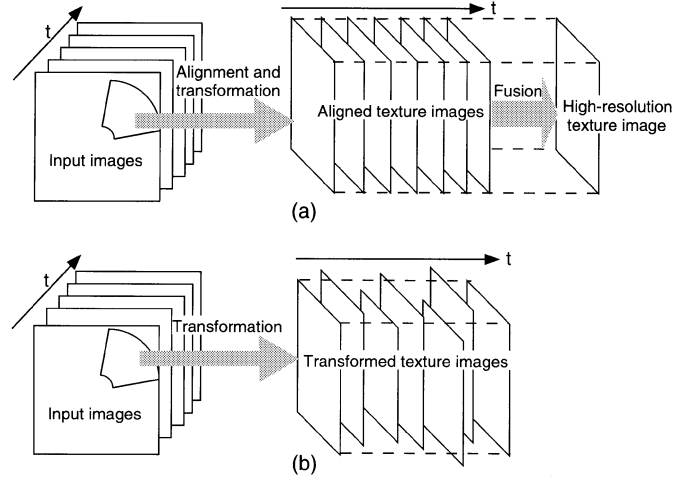Our method improves the resolution of each surface texture on the geometrical 3-D model by fusing observed textures in an image sequence. This method is based on the superresolution modeling concept.

As shown in Fig. 3(b), observed textures are not aligned precisely on the texture coordinates because the generated surface model has observational errors. To fuse these textures, observed textures should be aligned precisely. We precisely adjust the texture coordinates by using the template matching method shown in Fig. 3(a). Then, we estimate the high-resolution texture on 3-D model surface from omnidirectional input image sequence by superresolution and image mosaic. However, the resolution of input images depends on the position, height and posture of the object plane. Moreover, a catadioptrical omnidirectional image sensor has a large resolution difference corresponding to the depression angle. Since the high-resolution textures were optimized on the target plane on superresolution modeling, the input image resolutions also should be evaluated on the target plane for image registration of the resolution enhancement. Image resolution is not constant with the low-resolution part of the input image having a negative influence on the precise alignment by template matching. However, the high-resolution part is more important for the precise alignment.

Therefore, this method uses a resolution-weighting table that reflects the input resolutions on the target plane for image matching and image registration.

A. Resolution Weighting Table

Here, we define the texture coordinate $P_t(u, v)$ on the target plane (see Fig. 2). The relation between $P_t(u, v)$ and $P_w(X, Y, Z)$ can be represented by (4).

$$\tan\theta = \frac{Y}{X} = \frac{y}{x} \tag{1}$$

$$x = Xf \frac{b^2 - c^2}{(b^2 + c^2)Z - 2bc\sqrt{X^2 + Y^2 + Z^2}} \tag{2}$$

$$y = Yf \frac{b^2 - c^2}{(b^2 + c^2)Z - 2bc\sqrt{X^2 + Y^2 + Z^2}} \tag{3}$$

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = u \begin{bmatrix} u_X \\ u_Y \\ u_Z \end{bmatrix} + v \begin{bmatrix} v_X \\ v_Y \\ v_Z \end{bmatrix} + \begin{bmatrix} c_X \\ c_Y \\ c_Z \end{bmatrix} \tag{4}$$
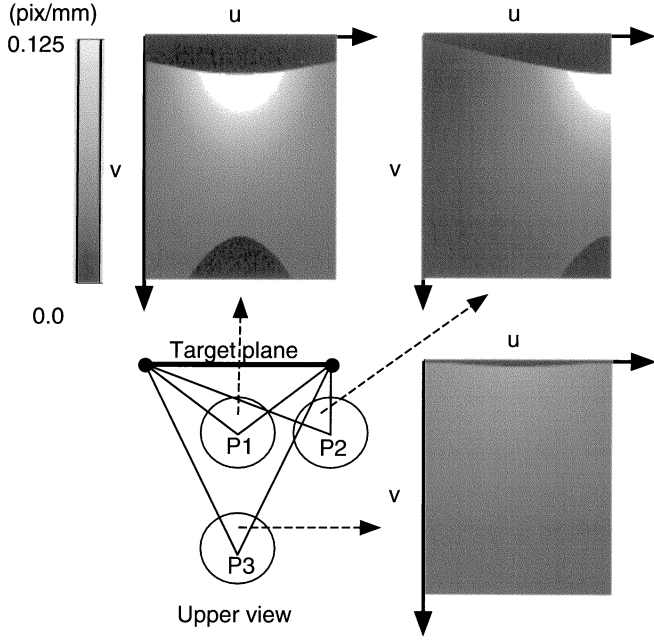
Fig. 4.   Resolution weighting table.

Fig. 5.   Search region for template matching.

$[u_X, u_Y, u_Z]^T$, $[v_X, v_Y, v_Z]^T$, and $[c_X, c_Y, c_Z]^T$ indicate unit vectors parallel to the u and v axes and the origin of the texture coordinate on the sensor coordinate, respectively. The resolution-weighting table is defined by the following equation:

$$W(u, v) = \frac{\sqrt{\Delta x^2 + \Delta y^2}}{\sqrt{\Delta u^2 + \Delta v^2}} \qquad (5)$$

Here, $(\Delta u, \Delta v)$ and $(\Delta x, \Delta y)$ are resolutions of texture and omnidirectional input image coordinates, respectively. The relation between $(\Delta x, \Delta y)$ and $(\Delta u, \Delta v)$ can be represented by the (2)–(4). We prepared the resolution-weighting table for each of the target planes at each sensor position. Fig. 4 shows examples of a resolution weighting table of a target plane at different sensor positions P1, P2, and P3. The top and the bottom of the target plane cannot be observed from the nearest sensor position P1. Black regions represent unobserved parts in the target plane. However, a high-resolution region (white region) also appears at the upper side of the target plane. On the other hand, from position P3, image resolution is not so high yet a full target plane can be observed. Moreover, the proposed weighting table contains the sensor resolution specification relative to the depression angle, because the table is derived from the projective relationship. Fig. 4 shows from the sensor resolution specification that the resolution of the upper side is higher than that of the lower.

B. Alignment of Texture Images

In this section, we describe how to align texture images by using template matching. As mentioned before, the input image resolutions are greatly changed relative to the position and the posture of target plane and the sensor resolution specification. The image resolution is not constant with the low-resolution part of the input image having a negative influence on the precise alignment by template matching. Our template matching method is weighted by the resolution-weighting table to
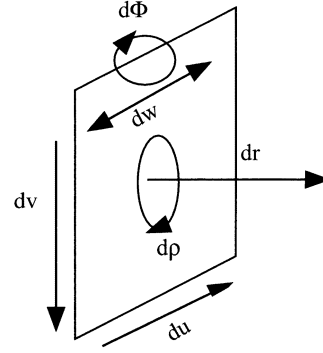
reflect the image resolution. Equation (6) shows the weighted correlation.

$$C = \sum_{u,v}^{N} W_{tr}(u,v) \frac{(f_t(u,v) - \bar{f}_t)(f_r(u,v) - \bar{f}_r)}{\sqrt{\sigma f_t \sigma f_r}}$$

$$\sigma f_t = \sum_{u,v}^{N} W_{tr}(u,v)(f_t(u,v) - \bar{f}_t)$$

$$\sigma f_r = \sum_{u,v}^{N} W_{tr}(u,v)(f_r(u,v) - \bar{f}_r)$$

$$\bar{f}_t = \frac{1}{N} \sum_{u,v}^{N} f_t(u,v)$$

$$\bar{f}_r = \frac{1}{N} \sum_{u,v}^{N} f_r(u,v)$$

$$W_{tr}(u,v) = W_t(u,v) + W_r(u,v) \qquad (6)$$

Here, $f_t(u,v)$ and $f_r(u,v)$ are the texture images at the t and the reference frames, respectively. $W_t(u,v)$ and $W_r(u,v)$ are the resolution weighting tables at the t and the reference frames, respectively, and N is the number of pixels of target texture image. The reference frame is the frame in which the average of the weighting table is the highest value. Search parameters for matching are illustrated in Fig. 5. $du, dv, dr, dw, d\phi$, and $d\rho$ are shifts along the u and v axes, the distance from the sensor, the width of the object plane, the rotation around the Z axis and the rotation around the normal direction of the target plane, respectively.

C. Superresolution Modeling and Image Registration by Resolution Weighting Table

The proposed method is based on concepts of superresolution modeling. To observe a full target plane, we stitch images captured at different sensor positions. As each image is observed at a different position, sub-pixel displacement occurs among the stitched images. The surface model and sensor positions are estimated by above geometrical modeling method. Therefore, the resolution of the texture image can be improved on the surface model plane by using superresolution techniques such as a back projection method.

First, we decide an image size for the high-resolution texture. The resolution-weighting table indicates the input resolution on texture plane. Comparing the averages of the weighting table
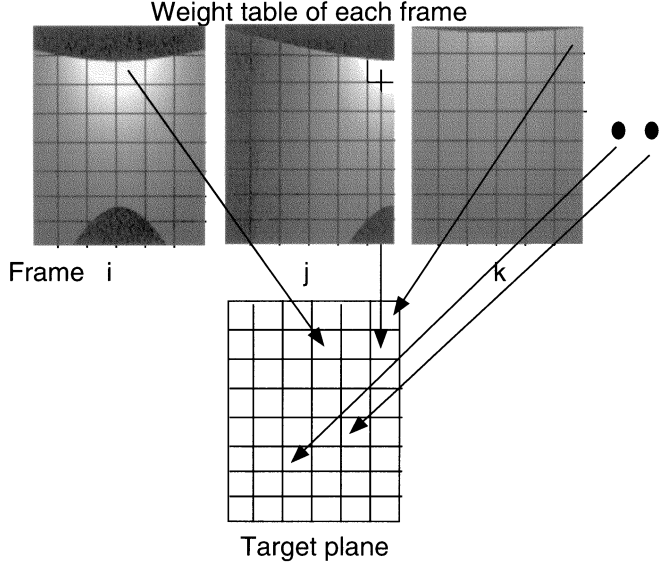
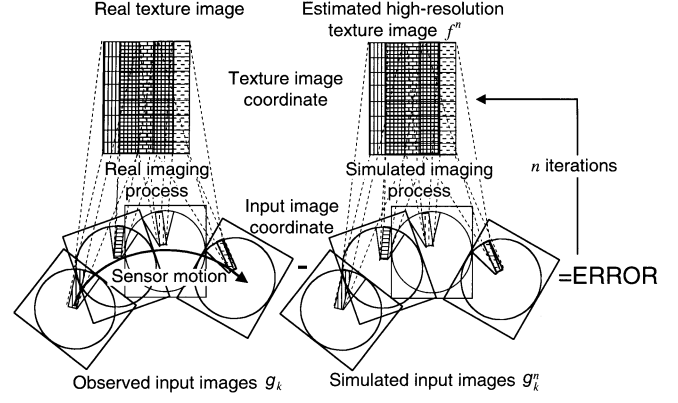Fig. 6. Selection of input frame based on resolution weighting table.

Fig. 7. Superresolution process by back projection method.

TABLE I
RESOLUTION IMPROVEMENT AGAINST THE MODELING ERROR

Standard deviation of locus error			PSNR
Model[mm]	Sensor[mm]	Pitching[deg.]	[dB]
0.0	0.0	0.0	44.5
50.0	15.0	0.0	40.4
100.0	30.0	0.0	40.7
200.0	60.0	0.0	19.7
100.0	30.0	2.0	36.0
Single input			25.0

of the whole texture plane at each frame, we use the highest average of the weighting to decide the image size for estimating high-resolution. The coefficient of resolution magnitude ratio multiplied by the average weight was estimated on preliminary experiments.

As shown in Fig. 6, the object plane is divided into a particular cell size. All resolution-weighting tables corresponding to each cell are compared. The magnitude of the weighting table and the frame number of the cell with the highest magnitude are memorized. If the magnitude of the weighting table so chosen is lower than a certain threshold, the highest of the remaining frames is selected. Until the sum of the selected resolution weights is higher than a particular threshold, other frames are selected. The selected frames are then used for superresolution. We decide the adequate image size for the high-resolution texture and register the input images at each cell to have uniform input image information through the target plane by this procedure. superresolution by back projection is done iteratively. The process is described in (8) and illustrated in Fig. 7.

$$g_k^n(\boldsymbol{P_i}) = \sum_{\boldsymbol{P_t}} f^n(\boldsymbol{P_t}) h^{\mathrm{PSF}}(\boldsymbol{P_t}) \qquad (7)$$

$$f^{n+1}(\boldsymbol{P_t}) = f^n(\boldsymbol{P_t}) + \sum_{P_i \in \cup_k P_{i\,k,P_t}} (g_k(\boldsymbol{P_i}) $$
$$ - g_k^n(\boldsymbol{P_i})) \frac{\left(h_{P_t P_i}^{\mathrm{BP}}\right)^2}{c \sum_{P_{i'} \in \cup_k P_{i\,k,P_i}} h_{P_t P_{i'}}^{\mathrm{BP}}} \qquad (8)$$

Where $\boldsymbol{P_t}$ and $\boldsymbol{P_i}$ are pixels on texture and input image coordinates, respectively. $f^n(\boldsymbol{P_t})$ is the estimated texture after n iteration on the texture coordinates. $g_k^n(\boldsymbol{P_i})$ is the predicted input image simulated from $f^n(\boldsymbol{P_t})$ as shown in (7). Here, h^{PSF} is the optical relation of HyperOmni Vision defined by the point spread function (PSF). The PSF and the resolution specification of the sensor were analyzed in [17], [18]. $g_k(\boldsymbol{P_i})$ is a real input image. As shown in (8), iterative estimation is done until the difference between the predicted input image and the real input image is minimized. $h_{P_t P_i'}^{\mathrm{BP}}$ is a back projection kernel that

means the contribution of $\boldsymbol{P_i}$ by $\boldsymbol{P_t}$. c is a constant normalizing factor. The kernel is calculated from (7) by a voting algorithm. $\boldsymbol{P_{i_{k,P_t}}}$ is the set of $\boldsymbol{P_i}$ that has an effect on $\boldsymbol{P_t}$. This superresolution method can simultaneously perform not only superresolution but also blur restoration.

VI. SIMULATION

The proposed method was evaluated by simulation. The size of the target plane was 2×2 m. The sensor was set in a 2×2 m square area at random. Distances between the target plane and the sensor were from 0.5 m to 2.5 m. The sensor height was 1.28 m. To numerically estimate the resolution improvement, we used the peak signal noise ratio (PSNR) between the resultant high-resolution texture and the ideal texture image.

Table I shows PSNR results comparing the proposed method from 25 input images with the bilinear interpolation from a single input. In this simulation, we assigned an estimated error to the sensor and the target edge positions. For example, (model 100, Sensor 30) indicates the situation of a 100 mm random error to the model position and a 30 mm random error to the sensor position. PSNRs were increased against the single input result in cases where the locus errors of the model and sensor were within 100 and 30 mm, respectively. The result was declined when the errors were 200 and 30 mm. For this reason the template matching failed because of its large locus errors. However, it is sufficient to apply to a real scene, because the locus errors were within 100 and 30 mm in experimental results [7]. Moreover, in practice the input images have a pitching noise when the images are captured by the robot. The fifth line in Table I
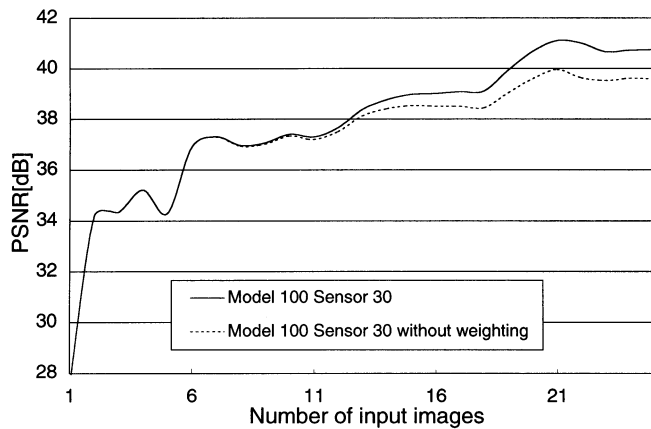
Fig. 8. Resolution improvement across the number of input images.

Fig. 9. Experimental environment.

Fig. 10. Result of geometrical surface model.

(a) Freme 00

(b) Frame 08

(c) Frame 11

(d) Frame 24

Fig. 11. Example of input images for resolution improvement and sensor motion.

shows the result where the simulated input images have a 2.0 deg pitching noise. The result was better than the single input case. From these results, we confirmed that our method could improve image resolution and was robust enough to estimated position error of both the sensor and the model.

The proposed method uses the weighting table for template matching and for the registering process. Fig. 8 shows the PSNR against a number of input images for resolution improvement. The solid and dotted lines in Fig. 8 indicate the results with and without using the weighting table, respectively. In a comparison of cases using and not using the weighting table, there are no differences if less than ten images are used. However, after about 20 frames, the PSNR converged at approximately 39.5 dB. This shows that the use of the weighting table is effective for improving resolution.

VII. EXPERIMENTAL RESULTS

An experiment was carried out using a real robot system (Nomad200) with HyperOmni Vision. The experimental environment was a corridor in a building as shown in Fig. 9. Three hundred and fifty input images were captured while a human operator controlled the robot. Edges not needed for geometrical
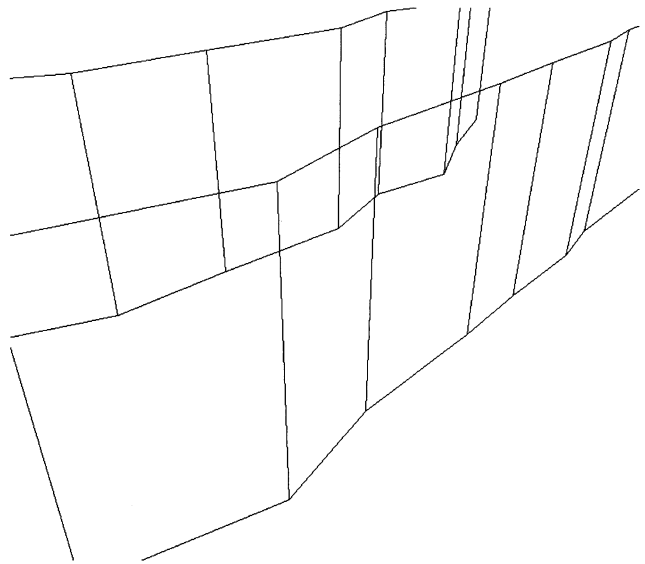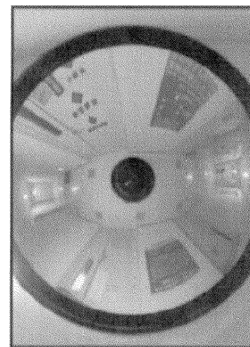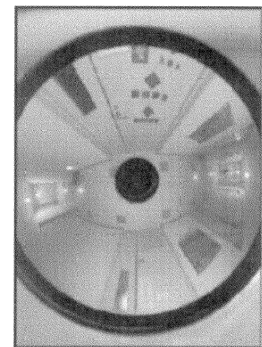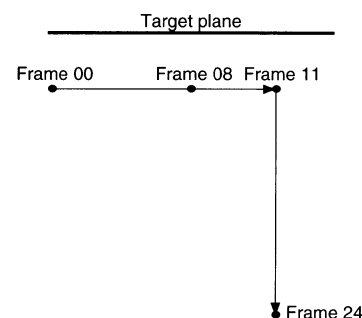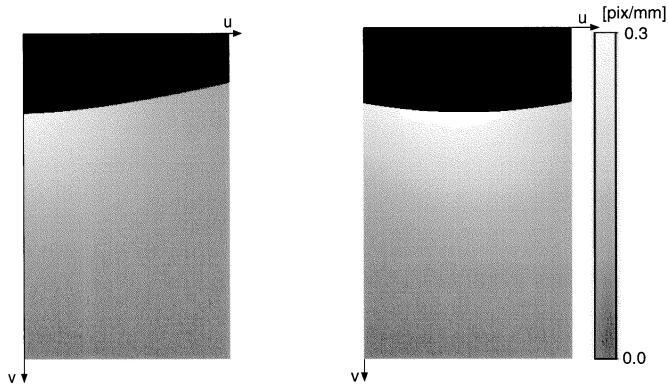
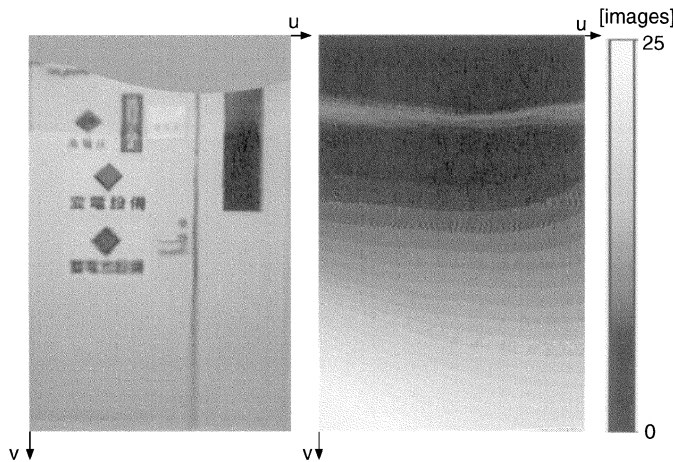Fig. 12. An example of resolution weighting table.

Fig. 13. Result of high-resolution texture and the number of images for resolution improvement (Left: High-resolution texture with proposed method, right: the number of registered images for resolution improvement).
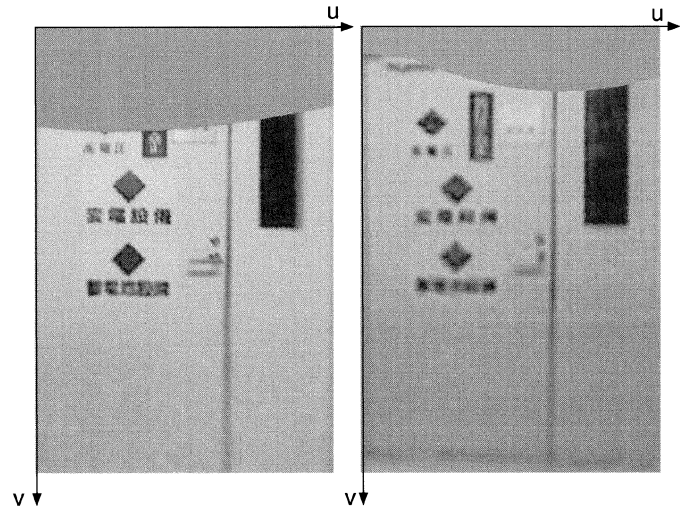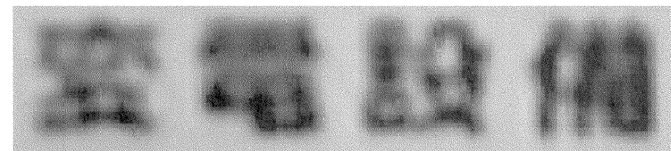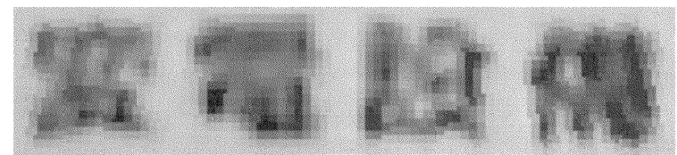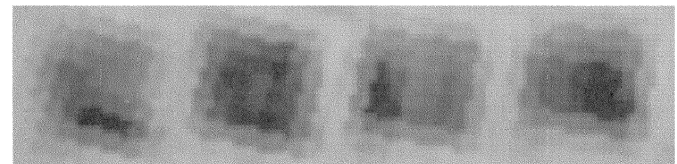

Fig. 14. Texture images transformed from a single image (Left: frame 05, right: frame 24).

(a) Proposed method

(b) Single input texture from frame 05

(c) Single input texture from frame 24

Fig. 15. Magnified image in part of texture.

modeling were eliminated by hand. Fig. 10 shows the results of the geometrical surface model.

The input images around the target plane were selected from the 350 input images. The maximum number of textures for superresolution was 25. Fig. 11(a)–(c) and (d) show the omnidirectional input images at frames 00, 08, 11, and 24, respectively. Each observed position relative to the target plane is shown in Fig. 11(e). Note that these input images were reversed owing to the optics of HyperOmni Vision being composed of a reflective mirror. Fig. 12 shows the resolution weighting tables at each frame.

As frame 00 was captured at the left side of the object plane, the weight at the left upper part was high. As frame 08 was captured at the middle front position of the object plane, the weight at the center upper part was high. From these results, the resolution weighting tables indicated the resolution difference corresponding to that part of the object plane and the sensor position relative to that plane.

Fig. 13 shows the results of the resolution improved image and also the number of input images used for resolution improvement. In this figure, many input images (white region) are used at the bottom of the output image. Two texture images in Fig. 14 were transformed from the input image at frames 05 and 24, respectively. Fig. 15(a)–(c) show a magnified image of Figs. 13 and 14, respectively. The texture image at frame 05 is

the image with the highest resolution of the input images. Frame 24 shows the widest field of view in the input images. From the aspect of the field of view, we can confirm the effect of the image mosaic because the widest field of view is maintained. Furthermore, it is interesting that the resolution of Fig. 15(a) is higher than that of Fig. 15(b), even if frame 05 has the highest weighted image on the focused area. The sampling artifact is decreased and the edge lines are smoother on the high-resolution image. This result shows that the effect of superresolution improves the resolution by fusing the information from multiple sub-pixel displaced images. Fig. 16(a) and (b) show results of the magnified high-resolution images with and without the weighting table for template matching and image registration, respectively. The edge lines are discontinuous and blurred

(a) Using the weighting table
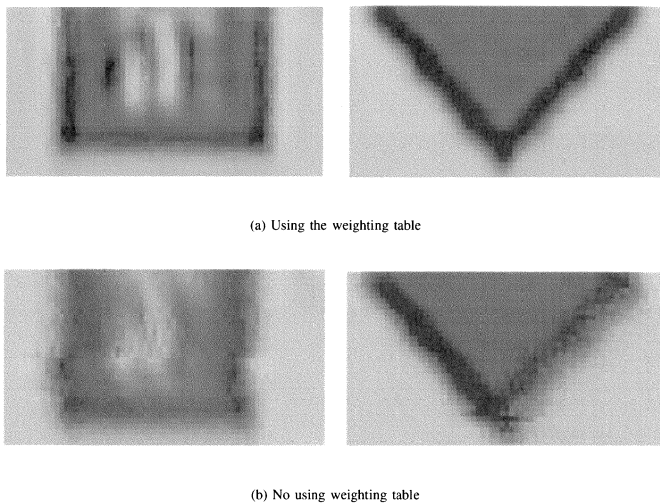

(b) No using weighting table

Fig. 16. Magnified high-resolution images using and no using the weighting table.

Fig. 17. Comparative result of virtual viewpoint image created from the generated 3-D model and textures (left: high-resolution texture, right: single input texture).

as shown in figures (b). The image contrast of (a) is higher than that of (b). These differences are caused by alignment errors of input images and also the use of unreliable low-resolution image parts. From these results, we confirmed the effectiveness of our resolution improving method.

However, the image brightness has changed at the horizon. This discontinuous part was caused by a changing in the light conditions while the robot is moving. This problem is one that will be addressed in the future.

On the left and right in Fig. 17 are high-resolution texture images by our proposed method and an ordinary texture transformed from a single input image, respectively. From the inputs, the position and direction of the sensor is different. These results also show that image resolution is improved.

VIII. CONCLUSION

An omnidirectional image sensor has an advantage that it can capture whole scenes around a sensor. It is useful for modeling because of its large field of view. However, the resolution of texture images is low due to the capture of this large field of view. Hence, resolution of the generated image is not sufficient for monitoring details. In this paper, we present new concepts of superresolution modeling and apply them to an omnidirectional image based geometrical modeling. We explained the distinction between superresolution imaging and superresolution modeling, and discussed that superresolution modeling is more suitable for application with geometrical modeling. The proposed method improves the resolution of textures mapped to a geometrical model by using superresolution modeling. High-resolution textures can be made from the input image sequences used for the geometrical modeling. Our resolution-weighting table was used for precise alignment of texture images, and was effective for improving resolution. We confirmed the improved resolution of texture images in both simulation and experimental results. The proposed resolution improvement method can also be used with other geometrical modeling methods, and with other omnidirectional image sensors with convex mirrors such as paraboloidal ones. In this paper, we show the prototype of a superresolution modeling system. However, each process needs to be studied more carefully and the algorithm optimized. For instance, we used a template marching scheme to align the input texture images, but a gradient descent alignment method would be more efficient for computational costs. What kind of matching method is most suitable for superresolution modeling needs considering. This is area we intend to address in the future.

REFERENCES

[1] Y. Yagi, "Omnidirectional sensing and its applications," *IEICE Trans. Inform. Syst.*, vol. E82-D, no. 3, pp. 568–579, 1999.

[2] K. Yamazawa, Y. Yagi, and M. Yachida, "New real-time omnidirectional image sensor with hyperboloidal mirror," in *Proc. 8th Scandinavian Conf. Image Analysis*, vol. 2, 1993, pp. 1381–1387.

[3] S. K. Nayer, "Catadioptric omnidirectional camera," *Proc. IEEE Computer Vision and Pattern Recognition*, pp. 428–488, 1997.

[4] Y. Yagi, Y. Nishizawa, and M. Yachida, "Map-based navigation for a mobile robot with omnidirectional image sensor COPIS," *IEEE Trans. Robot. Automat.*, vol. 11, pp. 634–648, Oct. 1995.

[5] L. Delahoche, C. Pegard, E. M. Mouaddib, and P. Vasseur, "Incremental map building for mobile robot navigation in an indoor environment," *Proc. of IEEE Int. Conf. Robotics Automation*, pp. 2560–2565, 1998.

[6] M. Etoh, T. Aoki, and K. Hata, "Estimation of structure and motion parameters for roaming robot that scans the space," in *Proc. Int. Conf. Computer Vision*, 1999, pp. 579–584.

[7] Y. Yagi, H. Hamada, N. Benson, and M. Yachida, "Generation of stationary environment map under unknown robot motion," *Proc. IEEE/IROS Int. Conf. Intelligent Robots Systems*, pp. 1487–1492, 2000.

[8] J. Simamura, N. Yokoya, H. Takemura, and K. Yamazawa, "Construction of an immersive mixed environment using an omnidirectional image sensor," *Proc. IEEE Wksp. Omnidirectional Vision*, pp. 62–69, 2000.

[9] H. Kawasaki, T. Yatabe, K. Ikeuchi, and M. Sakauchi, "Construction of 3-D city map using EPI analysis and DP matching," presented at the Proc. Asian Conf. Computer Vision, 2000, pp. 1149–1155.

[10] H. Kawasaki, K. Ikeuchi, and M. Sakauchi, "EPI analysis of omni-camera image," in *Proc. IAPR Int. Conf. Pattern Recognition*, vol. 1, 2000, pp. 379–383.

[11] R. A. Hicks, D. Pettey, K. Daniilidis, and R. Bajcsy, "Closed form solutions for reconstruction via complex analysis," *J. Math. Imag. Vis.*, vol. 13, no. 1, pp. 57–70, 2000.

[12] E. Brassart, L. Delahoche, C. Cauchois, C. Drocourt, C. Pegard, and M. Mouaddib, "Experimental results got with the omnidirectional vision sensor," *Proc. IEEE Workshop Omnidirectional Vision*, pp. 145–160, 2000.

[13] T. Svoboda, T. Pajdla, and V. Hlavac, "Epipolar geometry for panoramic cameras," in *Proc. European Conf. Computer Vision*, 1998, pp. 218–232.

[14] P. Chang and M. Hebert, "Omni-directional structure from motion," *Proc. IEEE Workshop Omnidirectional Vision*, pp. 127–133, 2000.

[15] K. Yamazawa, Y. Yagi, and M. Yachida, "3-D line segment reconstruction by using HyperOmni vision and omnidirectional Hough transforming," in *Proc. Int. Conf. Pattern Recognition*, vol. 3, 2000, pp. 487–490.

[16] S. K. Nayar and A. Karmarkar, "360 × 360 mosaics," *Proc. IEEE Computer Vision Pattern Recognition*, vol. II, pp. 388–395, 2000.

[17] H. Nagahara, Y. Yagi, and M. Yachida, "superresolution from an omnidirectional image sequence," *Proc. IEEE Int. Conf. Industrial Electronics, Control, Instrumentation*, pp. 2559–2564, 2000.

[18] ——, "Resolution improving method from multi-focal omnidirectional images," *Proc. IEEE Int. Conf. Image Processing*, pp. 654–657, 2001.

[19] T. Kawanishi, K. Yamazawa, H. Iwata, H. Takemura, and N. Yokoya, "Generation of high-resolution stereo panoramic images by omnidirectional imaging sensor using hexagonal mirrors," in *Proc. Int. Conf. Pattern Recognition*, vol. 1, 1998, pp. 485–489.

[20] R. Y. Tsai and T. S. Hang, "Multiframe image resolution and registration," *Adv. Comput. Vis. Image Process*, vol. 1, pp. 317–339, 1984.

[21] S. Pelg, D. Keren, and L. Schweitzer, "Improve image resolution using subpixel motion," *Pattern Recognit. Lett.*, pp. 233–226, 1987.

[22] M. Irani and S. Pelg, "Improving resolution by image registration," *Comput. Vis., Graph., Image Process.*, vol. 52, pp. 231–239, 1991.

[23] R. Schultz and R. Stevenson, "Extraction of high-resolution frames from video sequence," *IEEE Trans. Image Processing*, vol. 5, pp. 996–1011, June 1996.

[24] P. Cheeseman, B. Kanefsky, R. Kraft, J. Stutz, and R. Hanson, "superResolved Surface Reconstruction from Multiple Images," NASA Ames Research Center, Tech. Rep., FIA-94-12, 1994.

[25] S. Mann and R. W. Picard, "Virtual bellows: Constructing high quality stills from video," *Proc. IEEE Int. Conf. Image Processing*, pp. 363–367, 1994.

[26] D. Capel and A. Zisserman, "Automated mosaicing with superresolution zoom," *Proc. IEEE Int. Conf. Computer Vision Pattern Recognition*, pp. 885–891, 1998.

Yasushi Yagi received the B.E. and M.E. degrees in control engineering in 1983 and 1985, respectively, and the Ph.D. degree in 1991, all from Osaka University, Osaka, Japan.

In 1985, he joined the Product Development Laboratory, the Mitsubishi Electric Corporation, where he was working on robotics and inspections. In 1990, he was a Research Associate for information and computer science, Faculty of Engineering Science, Osaka University. From 1993 to 1996, he was Lecturer of systems engineering, Faculty of Engineering Science, Osaka University. In 1996, he was an Associate Professor of systems and human science, Graduate School of Engineering Science, Osaka University. From June 1995 to March 1996, he was an academic visitor of Department of Engineering Science, University of Oxford. His research interests include computer vision and robotics.

Hajime Nagahara received the B.E. and M.E. degrees in electrical and electronic engineering from Yamaguchi University, Ube, Japan in 1996 and 1998, respectively. He received the Ph.D. degree in system engineering from Osaka University, Toyonaka, Japan in 2001.

He is currently a Research Associate at Osaka University in the Japan Society for the Promotion of Science. His research interests include image processing, computer vision, and virtual reality.

Masahiko Yachida received the M.S. degree from Osaka University, Toyonaka, Japan, in 1971, and holds a D.Eng. degree.

He became a Research Associate at Osaka University. From 1967 to 1968, he was a Visiting Researcher at the Denmark Nuclear Power Laboratory, Denmark, The Netherlands. From 1973 to 1974, he was Research Associate at the University of Illinois, Urbana, and from 1980 to 1981, he was a Research Fellow at Hamburg University, Hamburg, Germany. In 1983, he was CDC professor at the University of Minnesota, Minneapolis. His research interests include computer vision, image processing, artificial intelligence, and mobile robots. He is the author of books on robot vision (Shokodo Press, Ogawa publication award), computer vision (Maruzen Press, Japan; editor), and others.

Dr. Yachida is a member of the Information Processing Society, the Robotics Society, the Artificial Intelligence Society, and others.