Title	Compound catadioptric stereo sensor for omnidirectional object detection
Author(s)	Sagawa, Ryusuke; Kurita, Naoki; Echigo, Tomio et al.
Citation	2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). 2004, 3, p. 2612-2617
Version Type	VoR
URL	https://hdl.handle.net/11094/14093
rights	c2004 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	

Compound Catadioptric Stereo Sensor for Omnidirectional Object Detection

Ryusuke Sagawa   Naoki Kurita   Tomio Echigo   Yasushi Yagi
The Institute of Scientific and Industrial Research,
Osaka University
8-1 Mihogaoka, Ibaraki, Osaka, 567-0047, JAPAN
Email: {sagawa,knaoki,echigo,yagi}@am.sanken.osaka-u.ac.jp

Abstract— This paper describes a novel system for detecting objects close to our sensor. For real time detection and portability, we have developed a small sensor with compound spherical mirrors. Since an object is projected onto each mirror, our method computes the range by a catadioptric stereo method. Our method creates a lookup table of corresponding points for an infinite range. If an object is close enough to the sensor, the projected points of the object are different from these corresponding points. Thus, our method can detect near objects by taking the differences in intensity of the corresponding points between the images in the mirrors. We show the experimental setup of our sensor and the result for detecting near objects.

I. INTRODUCTION

It is an important function for various applications to be able to detect near objects . For example, robots must detect approaching obstacles in order to avoid them. If people can detect a stranger stealing up behind them, they can avoid trouble before suffering damage. A system with a wide field-of-view and which is light-weight is needed for these applications.

In this paper, we describe a new sensor that we have developed for detecting objects, that are within a threshold range. Our new sensor consists of a single camera and multiple mirrors. with which we construct a catadioptric stereo image. Since the size of each mirrors is less than 1.5cm with a weight of only 3g, the system is very small and lightweight; and can be easily fastened to a person. Because the purpose of our system is to detect objects that are within a threshold range, our method determines if the object is far away or not, instead of searching corresponding points along epipolar lines. Our sensor can therefore detect objects in real time.

Several omnidirectional stereo systems have already been proposed. Two panoramic images acquired by rotating a camera, are used for omnidirectional stereo systems[1], [2]. Since images must be acquired one at a time in order to detect objects in real time, moving a camera is not suitable for this purpose. Catadioptric stereo methods with two cameras [3], [4] use parabolic and hyperbolic mirrors for each camera to generate an omnidirectional image. As a sensor fastened to a person should be small and light, two cameras are too large. A catadioptric stereo system with a single camera [5], [6] uses multiple mirrors to generate multiple virtual viewpoints, and computes distances by

epipolar geometry between those virtual cameras. Southwell et al.[6] proposed an omnidirectional stereo system using two mirrors with different curvatures. This approach is suitable for small sensors, so we developed a novel small mirror for our catadioptric stereo system. Since the compound spherical mirrors which we propose in this paper are very small, the accuracy of distance measurements is lower than that of a wide-baseline stereo system. Therefore, we propose a stereo system with multiple spherical mirrors to increase the robustness of computing corresponding points while keeping the size small; a narrow- and multi-baseline stereo system. We have developed a hyperbolic mirror for omnidirectional imaging [7]. However, because it is difficult to create small hyperbolic mirrors, our new sensor has spherical mirrors instead, since these are easier to make.

In the following sections, Section II describes the configuration of the mirrors, and Section III describes our new method for detecting objects with these mirrors. We show our experimental results in Section IV. Finally, we summarize our method in Section V.

II. OMNIDIRECTIONAL SENSOR WITH COMPOUND SPHERICAL MIRRORS

Our new omnidirectional sensor consists of compound spherical mirrors. Fig. 1 shows the experimental setup of the camera. The sensor has a large mirror and seven small mirrors, the diameters of which are 10mm and 4mm (see Fig. 2).

The rays from an object hit each mirror, and the reflected rays are projected onto the image plane. Fig. 3 is an example of the image of the sensor. Because the centers of the side mirrors are not on the camera axis, they are distorted. The rays which hit different mirrors, intersect each other at the object position, thus, we can compute the range of the object by triangulation. However, since the length of the baseline is the distance of points where the rays hit the mirrors, the baseline is quite narrow. It is difficult to compute the range of objects at an arbitrary range. Our method, therefore, only determines whether or not an object is at infinite range. This approach is applicable to a narrow baseline system and the computational cost is low.
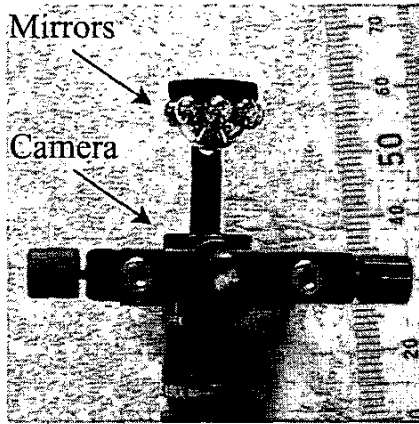
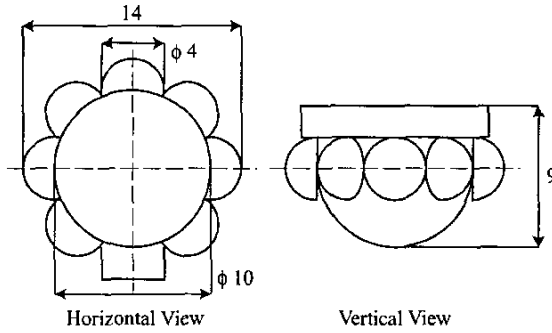Fig. 1. Omnidirectional camera with compound spherical mirrors.

Horizontal View   Vertical View

Fig. 2. Configuration of compound spherical mirrors.

III. DETECTION OF NEAR OBJECTS USING A CATADIOPTRIC STEREO SENSOR

Our new method detects near objects using a single stereo camera with multiple mirrors. Most stereo algorithms, such as [8], [9], search several disparities to find corresponding points of objects at different distances. However, it is difficult for our narrow-baseline stereo system to compute distances accurately. Our method searches for only one disparity to find the corresponding points, that is the disparity for infinite range; the disparity is zero for a parallel two-camera stereo system. If the object is sufficiently close to the camera, the colors of corresponding points for infinite range become different. Thus, our system can detect the object. Since we compare the image of the central mirror with that of multiple side mirrors, we increase the robustness of detecting the disparity with a multi-baseline stereo system.

The range of objects to be detected is determined by the length of the baseline and the image resolution. The length of baseline of our system is fixed because it is determined by the shape of the mirrors. Hence, we adjust the range to be detected by changing the image resolution.
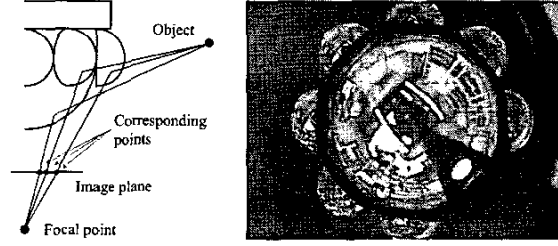

Fig. 3. An example of the sensor's image.

A. Computing Corresponding Points for Infinite Range

In this section, we compute the corresponding points when an object at infinite range is found in several mirrors of our sensor. Fig. 4 shows the situation for the central mirror. O is the origin of the camera coordinate system. d is the distance from the camera origin to the center of the mirror C, the radius of which is r. When the object is projected onto the point x in the central mirror, the ray from the object hits the mirror at the point p, $\angle pCO = \theta$, and $\angle pOC = \phi$. x is defined using γ,

$$(c_x + l\cos\gamma, c_y + l\sin\gamma) \tag{1}$$

in the image coordinate system, where (c_x, c_y) is the image center. Then,

$$\tan\phi = \frac{l}{f} \tag{2}$$

where f (in pixels) is the focal length of the camera. By considering the triangle $\triangle pCO$,

$$\tan\phi = \frac{r\sin\theta}{d - r\cos\theta}. \tag{3}$$

We compute θ by solving (3).

The incident angle α is $\theta + \phi$, because the normal vector at p is $\vec{Cp}$. Since the incident and reflected angles are the same,

$$\beta = \alpha + \theta = 2\theta + \phi. \tag{4}$$

Thus, the ray vector u from the mirror to the object becomes

$$(\cos\gamma\sin\beta, \sin\gamma\sin\beta, -\cos\beta). \tag{5}$$

If the object is at infinite range, the ray directions from the object to the mirrors are parallel. Fig. 5 shows the situation when the same object is found in one of the side mirrors. The incident ray vector is parallel to u. Since the unit vector v_0' from the origin O to the center of the side mirror C' is known,

$$\cos\beta' = -u \cdot v_0'. \tag{6}$$

In a manner similar to the case of the center mirror, we compute θ' by solving

$$\tan\phi' = \frac{r'\sin\theta'}{d' - r'\cos\theta'}, \tag{7}$$

where $\phi' = \beta' - 2\theta'$ and $d' = \sqrt{d^2 + r^2}$.
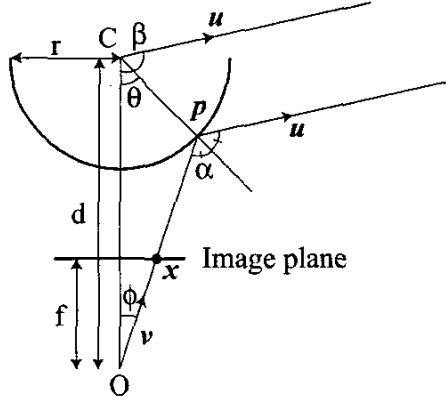
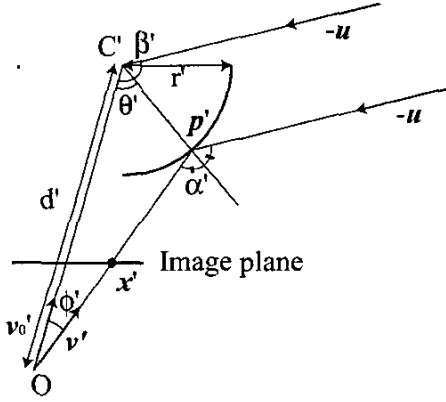Fig. 4. The ray direction reflected from the center mirror.

Fig. 5. The direction of the ray reflected from one of the side mirrors.

If the unit vector w is perpendicular to v'_0 and u, w is given by normalizing $v'_0 \times u$. Then, the vector $v' = (v'_x, v'_y, v'_z)$ from the origin to the point p', where the ray hits the side mirror, is computed by rotating v'_0 around w by angle ϕ'. Finally, the point x' at which the object is projected onto in the side mirror is computed as

$$(c_x + l' \cos\gamma', c_y + l' \sin\gamma'), \tag{8}$$

where

$$l' = f \frac{\sqrt{v'^2_x + v'^2_y}}{v'_z}, \tan\gamma' = \frac{v'_y}{v'_x} \tag{9}$$

Since the corresponding points are computed by preprocessing, our method refers to the lookup table to find the corresponding point when it computes the differences between the images produced by the mirrors.

B. Detecting Near Objects

If an object is close enough to the sensor, the projected point of the object is different from the corresponding point for infinite range. Our method, thus, takes the difference

in the intensity between the images from the mirrors. The intensity of the central mirror is $I(x)$, and the intensity of the corresponding point for a side mirror i is $I_i(x')$. The criterion $E(x)$ for detecting near objects becomes

$$E(x) = \frac{1}{N} \sum_i^N |I(x) - I_i(x')|, \tag{10}$$

where N is the number of side mirrors on which the object is projected. Then, we binarize $E(x)$ and make a histogram by counting the number of white pixels of the binarized image for each azimuth angle. If the near object is a person, the histogram can be used for detecting the person by using a human tracking technique, such as [10].

C. Adjusting Range of Detection

The range to be detected is the distance at which the disparity changes one pixel from that of infinite range. Thus, it can be computed by the relationship between the ray direction β and the projected point x. For simplicity, we assume the object is on the plane on which the centers of the central and one of the side mirrors are located. If an object is at a finite range as shown in Fig. 6, the derivative $\frac{\partial \beta}{\partial l''}$ can be computed from (7). This is the angle of the incident ray directions for the central and one of the side mirrors, when an object is detected. Since the length of baseline B is

$$B = u \cdot (p' - p), \tag{11}$$

the range to be detected R becomes

$$R = B \tan \frac{\partial \beta}{\partial l'}. \tag{12}$$

To adjust the threshold range, we change the image resolution. If the image size becomes half of the original, the threshold range R is computed as following using a new variable $l'' = \frac{1}{2}l'$:

$$R = \frac{1}{2} B \tan \frac{\partial \beta}{\partial l''}, \tag{13}$$

where B is assumed to be constant. Thus, the threshold range also becomes half of the original.

D. Calibrating the Sensor

To compute corresponding points, we require the distance d from the camera origin to the center of the central mirror and the focal length f. To calibrate the sensor, we first find the edges of the mirrors in the image. In Fig. 7, the distances from the image center to the edges are l_1 and l_2 for the central mirror and the side mirror. Then, the angle ϕ_1 and ϕ_2 becomes

$$\tan\phi_1 = \frac{l_1}{f}, \tan(\phi_1 + \phi_2) = \frac{l_2}{f}. \tag{14}$$

By considering the radius of the mirrors,

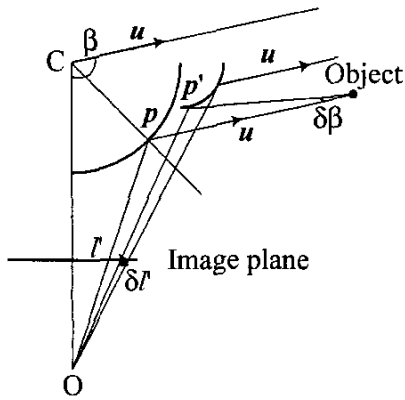$$\sin\phi_1 = \frac{r}{d}, \sin\phi_2 = \frac{r'}{\sqrt{d^2 + r^2}}. \tag{15}$$

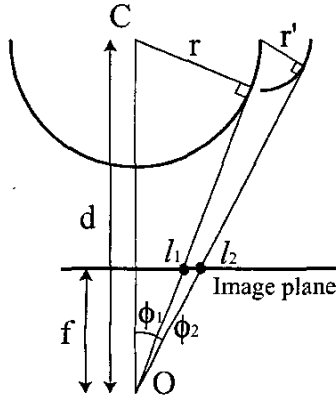Fig. 6. The difference in angle $\delta\beta$ if an object is in finite range.

Fig. 7. Calibration of sensor parameters d and f.

Hence, we solve the following equation for d by substituting $\tan\phi_1$ and $\tan(\phi_1 + \phi_2)$ using (15):

$$\frac{\tan(\phi_1 + \phi_2)}{\tan\phi_1} = \frac{l_2}{l_1}. \qquad (16)$$

We can compute ϕ_1 from (15). The focal length f is the given by (14).

IV. EXPERIMENT

In this section, we report on an experiment in detecting a person walking around our new sensor. Fig. 8 shows three examples of the result. The top row shows input camera images, The middle row shows $E(x)$ images, and the bottom row shows binarized images. In column (a), the person is far from the sensor. In column (b) and (c), he is approaching the sensor And is detected by it. The histograms of these binarized images are shown in Fig. 9. In this experiment, we empirically determined the threshold intensity was 25, where the range of intensity was 0-255. The image resolution was 640 × 480 pixels. Though our
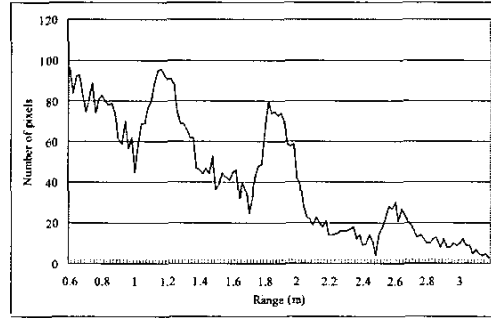

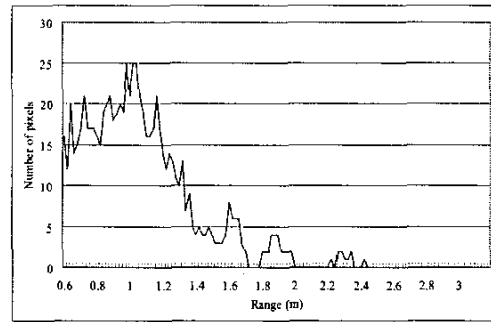Fig. 10. The number of detected pixels for each range.

Fig. 12. The number of detected pixels for each range by using the images at half resolution.

implementation has not been fully optimized, the computational time for each frame is 96(msec) using a Pentium4 3.0GHz processor. Fig. 10 shows the number of pixels which are detected after determining the threshold for each range. Though the number of detected pixels oscillates due to the swaying motion of the person, it typically decreases according to the person's range. Consequently, the range of detection in this experiment was about 3.2m.

Next, we change the image resolution to half the size of the original, 320 × 240 pixels. Fig. 11 shows the $E(x)$ and binarized images. Since the small disparities have disappeared due to reducing the image resolution, the number of detected pixels in the middle range is decreased. Fig. 12 shows the number of pixels which are detected after determining the threshold for each range. Since the reduction of image resolution works as a low-pass filter, the effect of the swaying motion is also reduced. Therefore, the range of detection for half resolution is shortened to about 1.8m.

V. CONCLUSION

This paper describes a new sensor for detecting near objects using a catadioptric stereo system with compound spherical mirrors. Since the sensor is small and light, it is easy to carry and fasten to a person. Our method takes the difference between corresponding points for infinite range.
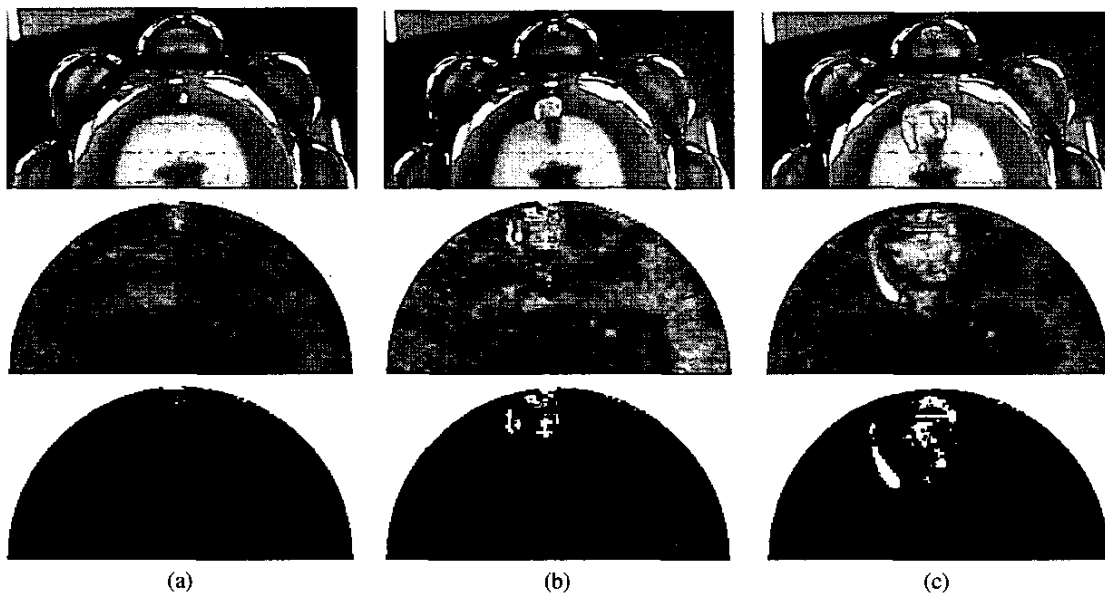
Fig. 8. Top: camera images, middle: $E(x)$ images, and bottom: binarized images. (a) far, (b) middle, and (c) near range.
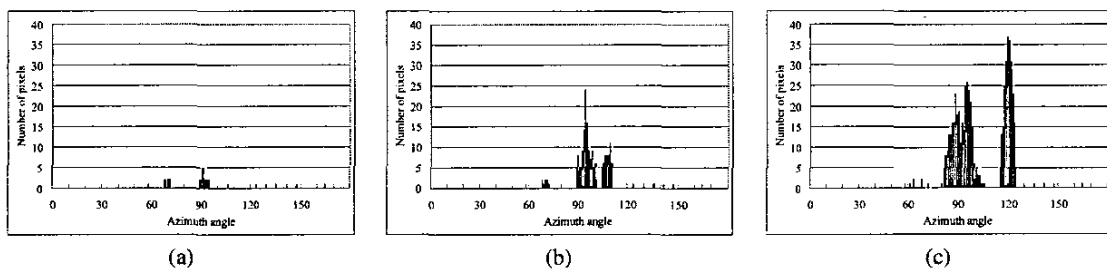

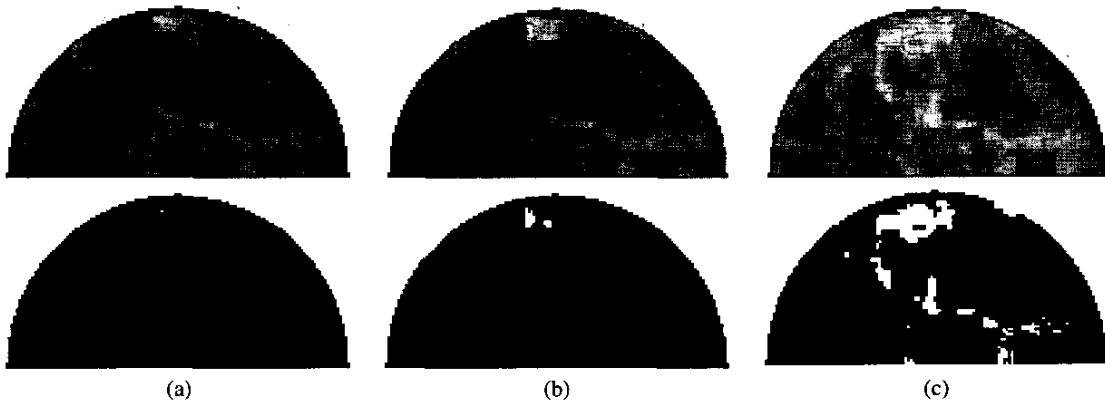Fig. 9. Histograms of the binarized images.

Fig. 11. Results of input images of half resolution. Top: $E(x)$ images, and bottom: binarized images. (a) far, (b) middle, and (c) near range.

If an object is close enough to the sensor, the projected points of the object are different from these corresponding points. Thus, our method detects near objects in real time without searching for exactly corresponding points. In future work, we will improve the shape of the mirrors to be more suitable for catadioptric stereo sensors.

REFERENCES

[1] H. Ishiguro, M. Yamamoto, and S. Tsuji, "Omni-directional stereo," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 14, no. 2, pp. 257–262. 1992.

[2] S. Peleg and M. Ben-Ezra, "Stereo panorama with a single camera," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, Ft. Collins, Colorado, June 1999, pp. 395–401.

[3] J. Gluckman and S. Nayar, "Real-time omnidirectional and panoramic stereo," in *Proc. of Image Understanding Workshop*, 1998.

[4] A. Chaen, K. Yamazawa, N. Yokoya, and H. Takemura, "Omnidirectional stereo vision using hyperomni vision," IEICE, Tech. Rep. 96-122, Feb. 1997, in Japanese.

[5] J. Gluckman and S. Nayar, "Rectified catadioptric stereo sensors," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, no. 2, pp. 224–236, 2002.

[6] D. Southwell, A. Basu, M. Fiala, and J. Reyda, "Panoramic stereo," in *Proc. International Conference on Pattern Recognition*, vol. A, 1996, pp. 378–382.

[7] K. Yamazawa, Y. Yagi, and M. Yachida, "Omnidirectional imaging with hyperboloidal projection," in *Proceedings IEEE/RSJ International Conference on Intelligent Robots and Systems*, vol. 2, July 1993, pp. 1029–1034.

[8] M. Okutomi and T. Kanade, "A multiple-baseline stereo," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 15, no. 4, pp. 353–363, 1993.

[9] O. Faugeras, B. Hots, H. Mathieu, T. Viéville, Z. Zhang, P. Fua, E. Théron, L. Moll, G. Berry, J. Vuillemin, P. Bertin, and C. Proy, "Real Time Correlation-Based Stereo: Algorithm. Implementations and Applications," INRIA, Technical Report N°2013, 1993.

[10] T. Mituyosi, Y. Yagi, and M. Yachida, "Real-time human feature acquisition and human tracking by omnidirectional image sensor," in *Proc. IEEE Conference on Multisensor Fusion and Integration for Intelligent Systems*, 2003, pp. 258–263.