

Development of Vision System on Humanoid Robot HRP-2

Yutaro Fukase

Institute of Technology, Shimizu Corporation, Japan
fukase@shimz.co.jp

Junichiro Maeda

Institute of Technology, Shimizu Corporation, Japan
jmaeda@shimz.co.jp

Yoshihiro Kawai

Intelligent Systems Institute, National Institute of Advanced Industrial Science and
Technology (AIST), Japan
y.kawai@aist.go.jp

Takashi Yoshimi

Intelligent Systems Institute, National Institute of Advanced Industrial Science and
Technology (AIST), Japan
tak-yoshimi@aist.go.jp

Fumiaki Tomita

Intelligent Systems Institute, National Institute of Advanced Industrial Science and
Technology (AIST), Japan
f.tomita@aist.go.jp

Abstract: Ministry of Economy, Trade and Industry of Japan had run an R&D project on humanoid robotics since 1998 for five years. In the first two years, platforms for humanoid robotics research were developed. In the next three years, researches were carried on for five different applications. We joined a group of the cooperative works by a human and a humanoid robot and took charge of a robot vision system. We designed a stereo camera module on a humanoid robot, developed some calibration methods and incorporated 3D image processing algorithm into the robot. Through this R&D, we showed a possibility to apply a humanoid robot for field of construction.

Keywords: humanoid robot, HRP-2, stereo vision, 3D image processing, VVV

1. INTRODUCTION

December 1996, Honda presented a humanoid robot "P2". This incident accelerated the research and development of humanoid robotics in Japan. Ministry of Economy, Trade and Industry of Japan had run an R&D project on humanoid robotics [1] since 1998 for five years. In the first years, platforms for humanoid robotics research were developed. In the next three years, researches were carried on for five different applications. The applications were maintenance tasks of industrial plants, building and home management services, human care services, operating industrial vehicles and cooperative works by a human and humanoid robot. We joined a group of the cooperative works by a human and a humanoid robot [2]. and took charge of a robot vision system.

We designed a stereo camera module on a humanoid robot, developed some calibration methods and incorporated 3D image processing algorithm into the robot.

2. TECHNICAL GOAL

The goal of the cooperative works by a human and humanoid robot was to build a terrapin hut together. Fig. 1 illustrates an image of the works. Two typical tasks were selected. One was to carry an external wall panel, and another was to mount the panel on the frame of a house. Some technologies were needed for achieving the following tasks.

- 1) Hardware/software to walk on uneven surfaces,

- and to prevent damage in the event of tipping
- 2) Environment recognition technology for working on a construction site
- 3) Communication between human and robot
- 4) Software to control arms and legs cooperatively

We took charge of “the environment recognition technology”. Developing the robot eyes for recognizing an environment and applying for detecting unevenness of a ground, searching a panel on the construction site and measuring the panel position and the grasping point on the panel.

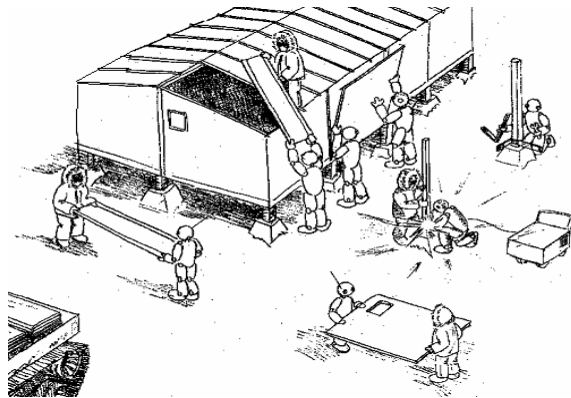
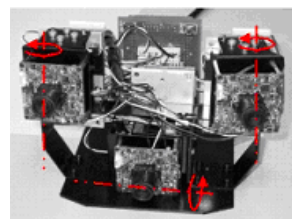


Fig. 1. Cooperative Works



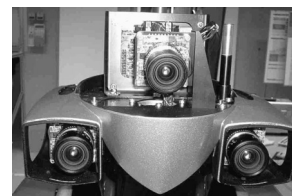
Fig. 2. HRP-2



a) Prototype System



b) System for HRP-2P



c) System for HRP-2

Fig. 3. Robot Vision

3. HRP-2

In order to achieve four technologies described above, a new humanoid robot, the name was HRP-2, had been developed. Fig. 2 shows the HRP-2. HRP-2's height is 1539mm and the weight is 58 kg including batteries. It has 30 *deg* of freedom (DOF) including two DOF for its hip. The cantilevered crotch joint allows the robot to walk in a confined area. Its highly compact electrical system packaging allows it to forgo the commonly used "backpack" used on other humanoid robots and to get up after falling. It has 6DOF force-torque sensors on each wrist. When HRP-2 carries a panel with a human operator, it senses force and torque in order to decide the walking direction and velocity.

4. VISION SYSTEM

HRP-2 has a trinocular stereo camera system on its head. The vision system was designed under restrictions which include the limit of weight, the limited space of the head and the measurement accuracy for the program. Versatile Volumetric Vision: VVV system [3], which had been developed by the 3D vision group of the AIST, was implemented for the HRP-2 image processing software.

4.1 Hardware

On the stages of developing the hardware, we made three types of system step-by-step. Fig. 3 shows the prototype, the system for HRP-2P and the system for HRP-2. The HRP-2P is a prototype model of the HRP-2.

1) Prototype Vision System

Resolution of each camera is 480x640 pixels and the intensity resolution is 256 grey level. Two horizontal cameras are separated by 120mm, and the third camera is located 60mm below the two horizontal cameras. Three cameras are used in order to increase 3D measurement accuracy and reduce stereo correspondence errors. A Stepping motors behind each camera changes the convergence according to the distance from the cameras to an object. In order to examine performances of the prototype system, it was mounted on HRP-1, which is based on Honda humanoid robot P-2, and captured several scenes. Fig. 4 shows the prototype vision system mounted on HRP-1. The experiments revealed that the cameras rotation mechanisms were vibrated due to a vibration generated from the motors of the robot joints

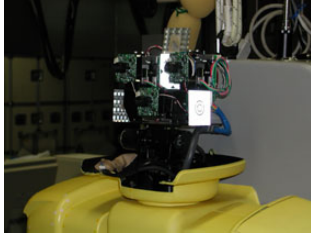


Fig. 4. Prototype Vision System on HRP-1

2) System for HRP-2P/HRP-2

For requirement of the limited weight, we could not use the motors that generate enough torque to stop the vibration of the camera's rotation mechanisms. Therefore a wide-angle C-mount lenses having a single focal length (8mm) was used to capture distant and nearby scenes simultaneously. The camera's shutter speeds were variable and adjusted automatically under several lighting situations. The total weight is less than 700g.

We confirmed performances of the vision system with the system for HRP-2P and designed the last system for HRP-2. The main differences between the system for HRP-2P and the system for HRP-2 are the length of the cameras base-line (120mm for HRP-2P, 140mm for HRP-2) and the appearance of the shields.

4.2 Software

The software for controlling HRP-2 was developed with Open Architecture Humanoid Robotics Platform (OpenHRP) [4]. The OpenHRP is a software platform for humanoid robotics and has been developed since previous term of the HRP. The OpenHRP consists of several modules to control body, to recognize voice, to generate voices and etc, and a task manager which integrates those modules. With this structure, each module can be developed independently and integrated to the software for controlling robot easily.

Stereo vision module was developed for one of the modules of OpenHRP. Each module of OpenHRP can be executed in a computer on a TCP/IP network. In order to connect the stereo vision camera on the robot's head, the stereo vision module was implemented in a computer inside HRP-2. When body control module sends a command to the stereo vision module, it carries out image processing and sends back a result.

5. DISTORTION CORRECTION

The captured images with HRP-2 were distorted due to the shields in front of the stereo vision system, and the wide-angle lenses. Since these distortions had adverse effects on image processing, we had to develop two distortion correction methods [5].

5.1 Shield Distortion

The cameras were covered by shields as shown in Fig. 5. It was difficult to derive a distortion model from relationship between a shield shape and a camera position deductively. Therefore we used a reference table which converts distorted image to undistorted one (Fig. 6). First, the camera without shield captures a panel, on which several circles were located periodically. Next the camera with a shield captures the same panel. For all circles in the each image, the positions of a center of gravity are calculated. Using a large circle near the image center as a reference, we can build correspondence of each circle's position between the distorted image and the undistorted image. Transformation for all pixels is performed by interpolation based on the correspondence of all circles.

Fig. 7(a) and Fig. 7(b) are captured images with and without a shield, respectively. Fig. 7(c) is a converted image of Fig. 7(a). Fig. 7(d) shows differences between Fig. 7(a) and Fig. 7(b). Fig. 7(e) shows differences between Fig. 7(b) and Fig. 7(c). In Fig. 7(d), influence of the distortion can be observed as white parts left side of circles. In Fig. 7(e) the white parts are decreasing.

Captured image with a shield becomes too dark to execute image processing. We take longer exposure time for the image with a shield. Therefore in Fig. 6 and Fig.7, the image with a shield is brighter than it without a shield. And circle patterns in Fig. 7(d) and 7(e) is the result of the difference of the exposure time between the images with and without shield. It is independent to the distortion for a shield.



Fig. 5. With and Without Shield

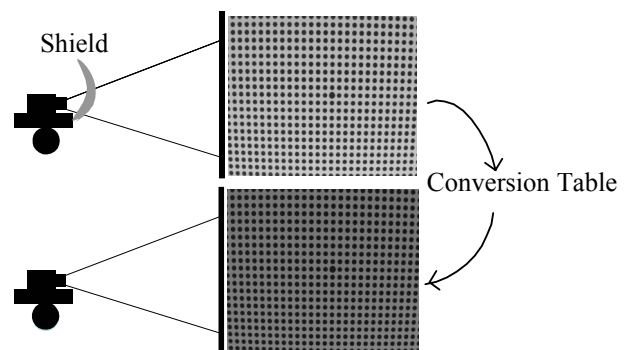


Fig. 6. Distortion for Shield

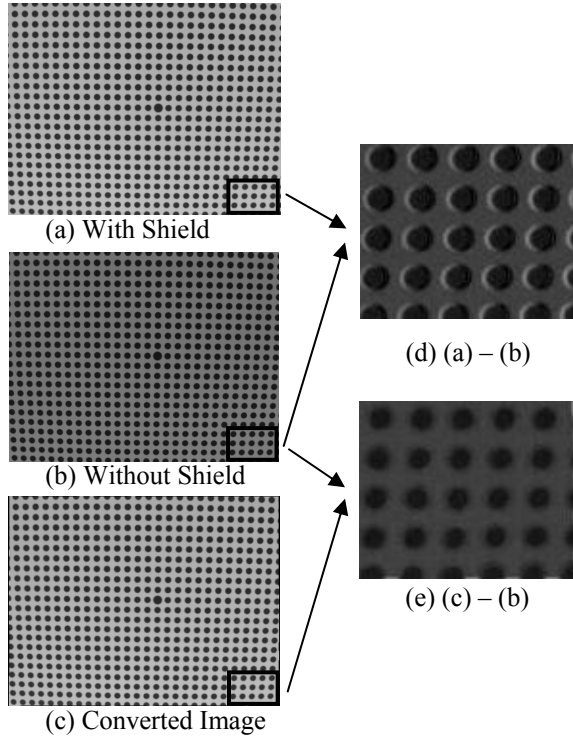


Fig. 7. Correcting Distortion for Shield

5.2 Lens Distortion

Fig. 8 shows aberration, which is caused by wide-angle lens. Relationship between distance from a principal point to a observed point r' and distance from the principal point to a ideal point r is below.

$$r' = r + k_1 \cdot r^3 + k_2 \cdot r^5 \quad (1)$$

Parameters, k_1 and k_2 , depend on the lens. Usually one coefficient k_1 is sufficient to correct the distortion, however, more coefficients are need for shorter focal length.

In order to correct this distortion, we also use the same image, which is captured for correcting the distortion for the shield. The lattice position \mathbf{P} on a plane is expressed as follows:

$$\mathbf{P} = i\mathbf{u} + j\mathbf{v} + \mathbf{w} \quad (2)$$

where i, j are integers. \mathbf{u}, \mathbf{v} are non-parallel vectors and \mathbf{w} is the origin. The position of the large circle is defined as \mathbf{w} . Vectors for column and row directions are \mathbf{u}, \mathbf{v} . Any point can be expressed by the parameters (i, j) . Data near the image center is generally only slightly distorted for the wide-angle lens, therefore the ideal positions without distortion for all lattice points are calculated using these data. Assigning the ideal positions and the observed positions into Eq. (1) and applying least-square method, we calculate the k_1 and k_2 . Fig. 9(a) shows a base image with lens distortion and Fig. 9(b) shows a

corrected image using the lens model. Distortion at rim of the image is corrected and the error was less than 0.2pixels.

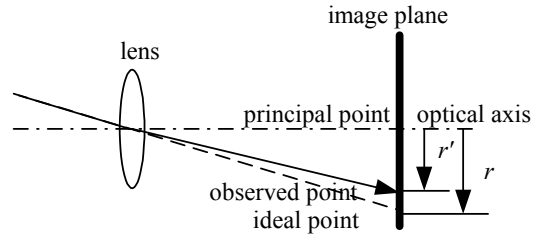


Fig. 8. Distortion for Lens

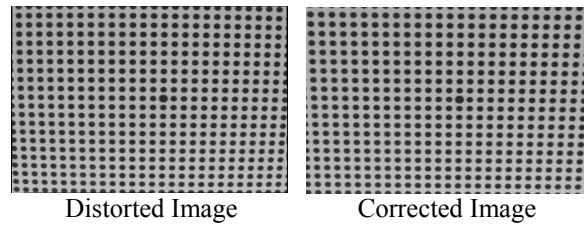


Fig. 9. Correction Distortion for Lens

6. DEMONSTRATION

Carrying a panel cooperatively with a human operator and putting the panel to build a wall of a terrapin hut were the tasks, in which developed technologies of our group were demonstrated. The vision system was used for autonomous tasks in former part of the demonstration.

6.1 Measuring Ground Displacement

HRP-2 could not take its step on 2cm unevenness of the ground. In order to avoid a rough ground, HRP-2 has to detect the range data of the ground [6]. Fig. 10 shows configuration of the robot detecting the rough ground. HRP-2 bends down its head and observes trapezoid area on the ground. In order to obtain the range data of the ground surface, a correlation stereo method is applied to two images captured by right and left cameras. Using a LINUX PC (Pentium3 1.26GHz CPU with 256MB memory) with acceleration technique proposed by Okada et al. [7], image processing for two VGA size, 256 grey level images could be executed within 4 seconds. Fig. 11 shows the stereo images and the obtained range data.

The ground displacement measurement error due to the discrete value of the disparity is approximately 20mm, but with interpolation it can be reduced to approximately 15mm. Fig. 12 shows histograms of z (vertical direction) value for a flat floor. Two curves are histograms with and without interpolation, which shows improvement by interpolation clearly. Because the origin of the foot coordinate is fixed on the ankle joint, which is

100mm above the floor, it has a peak at -100mm.

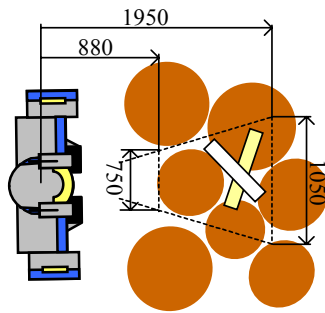
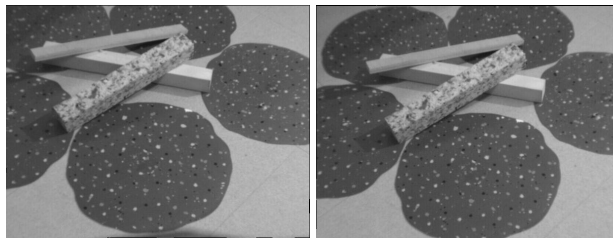
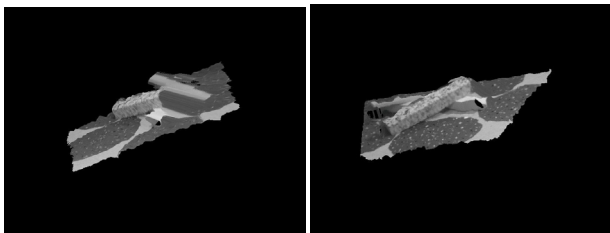


Fig. 10. Watching Uneven Ground



Left Image

Right Image



Reconstructed 3D Range Data

Fig. 11. Ground Displacement Measurement

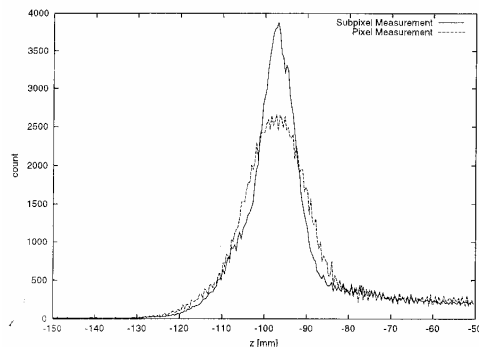


Fig. 12. Measurement Error

6.2 Approaching and Grasping Panel

1) Task Sequences

HRP-2 autonomously searches and approaches a storage site of panels, and grasps a panel end (Fig13). Following is the detailed sequence:

(a) A human operator gives “Start work” voice command to the HRP-2. HRP-2 turns its head to the

expected direction of a storage site of panels. The robot captures images and tries to detect the shape of the top of the pile of panels in the scene. When no panel is detected in the scene, it turns the head 10deg each and captures images again until the panel is finally detected. When the panel is detected, in order to confirm the result, the robot turns its head straightly to the direction and detects the panel once more. According to the obtained panel position and attitude, HRP-2 calculates the path to the first approaching point, which is 600m far from the end of the panel.

(b) HRP-2 walks the first approaching point. The robot turns its head to the direction of the panel and detects it again. According to the result, the robot walks in front of the panel end.

(c) HRP-2 bends its waist and neck joints and looks up the right near corner of the panel. The robot detects the corner and calculates position and attitude of the grasping point of the panel.

(d) In order to grasp the panel, HRP-2 turns its waist and moves its hands to the adequate position. Then the robot and the operator carry the panel together.

In this sequence HRP-2 compensates a walk error and a vision measurement error. HRP-2 can stably approach the storage of panels and grasp the panel.

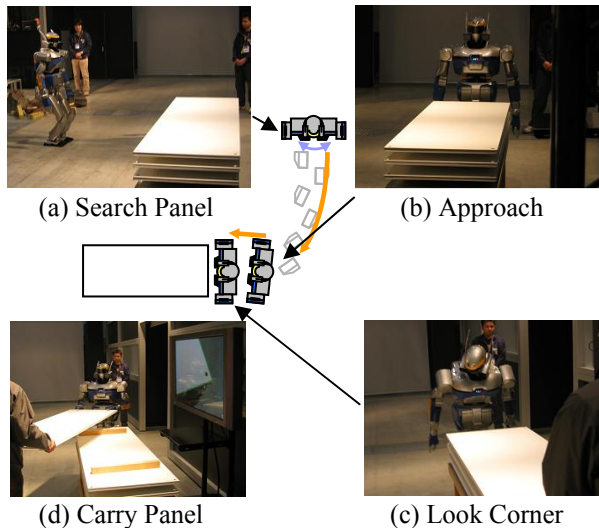


Fig. 13. Task Sequences

2) Detecting the Panel and its Corner

A segment-based stereo method [8] in the VVV system is applied to recognize a panel and a corner of the panel. The procedure of the panel recognition at the starting point of the work is as following:

- (a) Capturing stereo images simultaneously
- (b) Extracting boundary line from the images
- (c) Reconstructing 3D structure using segment-based stereo
- (d) Matching the reconstructed 3D structure with a shape model of the panel

When HRP-2 detects the corner of the panel in front of the panel end, the image processing is the same in principle. The grasping point was detected to within an error of 2mm and stably carrying a panel was accomplished.

7. SUMMARY

This presentation describes a vision system on a humanoid robot applied to constructing a terrapin hut. In response to design of the robot, a lightweight trinocular stereo camera system having auto shutter speed control was developed. An image processing software was implemented on OpenHRP, which was a software platform for humanoid robots. Calibration methods for distortion corrections were developed. Finally we demonstrated measuring a ground displacement, approaching a storage site of panels and grasping a panel.

ACKNOWLEDGMENTS

We express our thanks to the stuffs in AIST, Kawada Industries, Inc. and Yaskawa Electric Co., who cooperated with this R&D.

REFERENCES

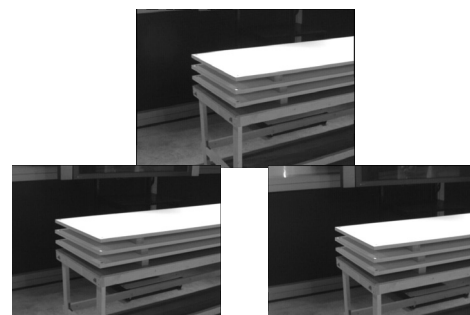
- [1] Hirukawa, "Overview of Humanoid Robotics Project of METI", *Proc. the 32nd Int. Symposium on Robotics*, pp.1478-1482, 2001.
- [2] Yokoyama, Maeda, Isozumi and Kaneko, "Application of Humanoid Robots for Cooperative Tasks in the Outdoors", *IEEE/RSJ IROS Workshop on Explorations towards Humanoid Robot Applications*, 2001.
- [3] Tomita, Yoshimi, Ueshiba, Kawai, Sumi, Matsushita, Ichimura, Sugimoto and Ishiyama, "R&D of Versatile 3D Vision System VVV", *Proc. IEEE Int'l Conf. on SMC*, pp.4510-4516, 1998.10
- [4] Kanehiro, Fujiwara, Kajita, K.Yokoi, Kaneko, Hirukawa, Nakamura, and Yamane, "Open architecture humanoid robotics platform", *Proc. the 2002 IEEE International Conference on Robotics & Automation*, pp.24-30, 2002.
- [5] Kawai, Tomita, Hirukawa and Fukase, "A Stereo Vision For The HRP-2 Humanoid Robot To Act Autonomously", *Proc. the 6th Asian Conference on*

Computer Vision, Vol.2, pp.745-759, 2004

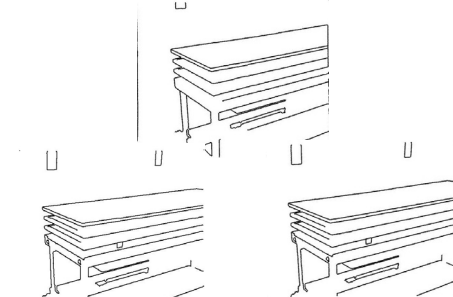
[6] Yoshimi, Kawai, Fukase, Araki and Tomita, "Measurement of Ground Surface Displacement using Stereo Vision and Mechanical Sensors on Humanoid Robots", *Proc. IEEE Conference on Multisensor Fusion and Integration for Intelligent Systems*, 2003

[7] Okada, Kagami, Inaba and Inoue, "Onbody Realtime Disparity Image and Flow Generation System based on Highspeed Correspondence Matching using PC", *Journal of the Robotics Society of Japan*, 18, 6, pp.896-901, 2000

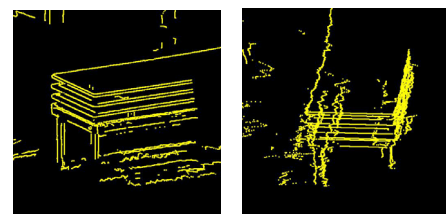
[8] Sumi, Kawai, Yoshimi and Tomita, "3D Object Recognition in Cluttered Environments by Segment-Based Stereo Vision", *International Journal of Computer Vision*, 46, 1, pp.5-23, 2002



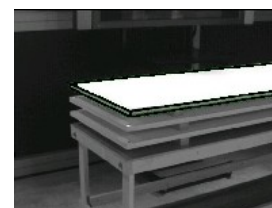
(a) Capturing Images



(b) Extracting Edge Lines



(c) Reconstructing Range Data



(d) Matching with a Model

Fig. 14. Detecting Panel