

Wheelchair Support by a Humanoid Through Integrating Environment Recognition, Whole-body Control and Human-Interface Behind the User

Shunichi Nozawa, Toshiaki Maki, Mitsuharu Kojima, Shigeru Kanzaki, Kei Okada, Masayuki Inaba

Abstract—In this paper, we treat with wheelchair support by a life-sized humanoid robot. It is quite essential to integrate whole-body motion, recognition of environment and human-interface behind the user in order to achieve this task. Contributions of this paper is whole-body control including pushing motion using the offset of the ZMP and observation of the attitude outlier, recognition of the wheelchair using particle filter and human-interface behind the person using face detection and recognition of gesture.

I. INTRODUCTION

It is important for humanoid robots expected to assist human daily life and added to various functions [1]–[3] to be capable of accomplishing nursing care tasks. Presently nursing care tasks require comprehensive support. In nursing care tasks, elemental techniques must be integrated [4] each other.

This paper deals with wheelchair support, including support for heavy works and interaction with people that are common elements of nursing care tasks. Here, we select such principal element as follows from among tasks in a wheelchair support scenario: (1) whole-body motion for transfer of a wheelchair (2) recognition of environment (3) interaction behind a person on the wheelchair. This paper



Fig. 1. Wheelchair transfer experiment outdoors

aims at the support that the humanoid robot pushes the wheelchair and the user can move to anywhere only by instructing where to go (Fig. 1).

S. Nozawa, T. Maki, M. Kojima, S. Kanzaki, K. Okada and M. Inaba are with Department of Mechano-Informatics, Graduate School of Information Science and Technology, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan {nozawa, maki, kojima, kanzaki, okada, inaba}@jsk.t.u-tokyo.ac.jp

II. BEHAVIOR INTEGRATED WITH SENSOR INFORMATION IN WHEELCHAIR SUPPORT

Almost all behavior of humanoid robots are integrated with sensor information in various levels in a broad sense.

In wheelchair support, the robot must become adapted to the environment. It is necessary for the robot to control the wheelchair adequately and to detect obstacles.

In addition, the robot should interact with the user. At the very least, it is preferable that the robot can push the wheelchair to desirable directions and suspend walking when the person hopes to stop walking.

In this paper, we integrate sensor information with behavior as follows (Fig. 2). The robot starts pushing the wheelchair

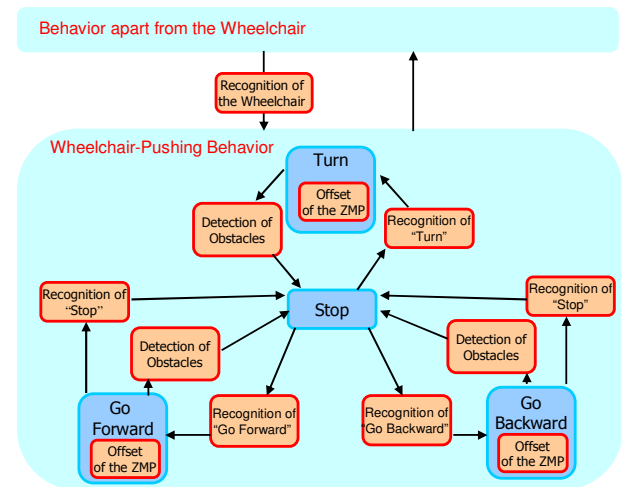


Fig. 2. Behavior Integrated with Sensor Information in Wheelchair Support

when it sees the person's following instructs by recognizing the gesture and detecting the face: "Go Forward", "Go Backward" and "Turn". While pushing the wheelchair, the robot observes the obstacles using somatic or visual sensors and attends to the person's "Stop" gesture. If the obstacles of "Stop" gesture are detected, the robot suspends walking. When the robot behaves apart from the wheelchair and returns to there, it self-localizes by recognizing the wheelchair.

III. WHOLE-BODY MOTION IN WHEELCHAIR SUPPORT

In this paper, the humanoid robot obtains the pushing force by offsetting the ZMP (Zero Moment Point). Furthermore, suppose the robot keeps on pushing when the wheelchair

would not move, motors of the robot may suffer from overload or the humanoid robot could possibly tumble down. Thus, the robot has to interpret whether the wheelchair moved or not, according to the sensory information.

A. Determination of ZMP-offset Appropriate to Hand External Forces

Fig.3 shows the Cart-Table-Model [5], the dynamic model which the humanoid robot obeying while pushing the wheelchair. We define the COG(Center Of Gravity), the

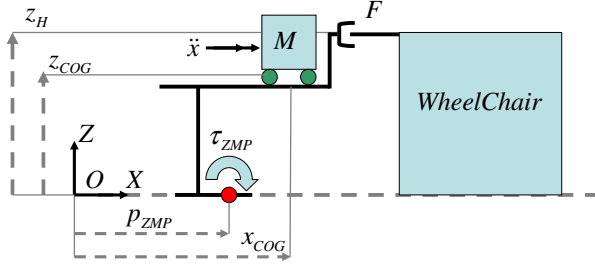


Fig. 3. The cart-table model affected by hand external forces

ZMP, the position of hand by relating the absolute coordinate O . Let x_{COG} be the direction of the wheelchair motion. z_{COG} is the height of the COG and \ddot{x} is the acceleration of the COG. F is the x-component of the reflected forces working on the humanoid hand. Equation 1 introduces p_{ZMP} as the ZMP.

$$\tau_{ZMP} = Mg(x_{COG} - p_{ZMP}) + M\ddot{x}z_{COG} + z_H F \quad (1)$$

Solving for p_{ZMP} results in Equation 2 .

$$p_{ZMP} = x_{COG} - \ddot{x} \frac{z_{COG}}{g} + \frac{z_H F}{Mg} \quad (2)$$

On one hand the third term in Equation 2 disappears when the robot does not push the wheelchair, On the other hand the robot and the wheelchair system obeys Equation 2 while pushing the wheelchair. This implies to add the offset to the ZMP enables the robot to push the wheelchair.

B. Obstacle Detection According to Estimation of Attitude

In wheelchair support, the robot can detect obstacles by the following two ways; the robot can recognize a visible obstacle using visual sensors, or by detecting the collision between the wheelchair and hindrances at a dead angle by utilizing somatic sensors.

In this section, we describe the latter method.

In following we managed the pitch-component of the robot's attitude (θ_{att}), for the pitch-component is the principal cause of tumble while transferring the wheelchair. When θ_{att} is over threshold, the robot detects collision with an obstacle.

This section introduces the method to learn threshold. In case that threshold is determined off-line, it is necessary to configure the threshold every time when the robot's behavior change.

Therefore, we modified the method to decide the threshold. The outline of the method is follows:

- (α) Learn parameters of attitude distribution for a period of time after the robot starts walking.
- (β) After the outlier of attitude is detected, stop walking.

(α) : In order to estimate parameter of attitude distribution, we employ the SDEM algorithm (Sequentially Discounting Expectation and Maximizing) [6]. The SDEM algorithm is adopted for calculation of distributive parameters in Gauss mixture model and has the feature being capability of incremental measurement.

The summary of the method to determine threshold employing the SDEM algorithm is Fig.4.

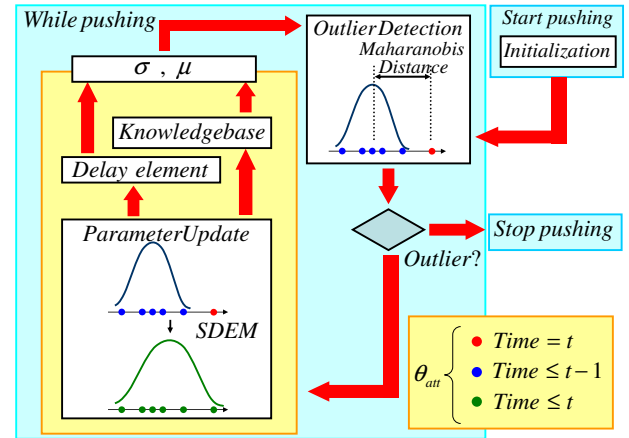


Fig. 4. Concept of threshold-decider using the SDEM and the Maharanobis distance

In this context, the SDEM algorithm is simplified as follows:

- 1) Definition

$$p(\theta_{att} | \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}(\theta_{att} - \mu)^2\right) \quad (3)$$

- 2) Initialization

$$\text{Set } \mu^{(0)}, \sigma^{(0)}, \bar{\sigma}^{(0)}, t := 1 \quad (4)$$

- 3) Parameter update Read θ_{att} from sensors and calculate

$$\mu^{(t)} = (1 - r)\mu^{(t-1)} + r\theta_{att}^{(t)} \quad (5)$$

$$\bar{\sigma}^{(t)} = (1 - r)\bar{\sigma}^{(t-1)} + r\theta_{att}^{(t)2} \quad (6)$$

$$\sigma^{(t)} = \bar{\sigma}^{(t)} - \mu^{(t)2} \quad (7)$$

(β) : The robot assumes the Maharanobis distance for the criterion of whether or not the current attitude is outlier, after parameters of attitude distribution have been calculated.

$$\text{Maharanobis distance} = \frac{|\theta_{att}^{(t)} - \mu^{(t-s)}|}{\sigma^{(t-s)}} \quad (8)$$

r , which is the parameter for forgetting out-of-date statistics, is decided experimentally on account of the period of walking. We use delay element in calculating the Maharanobis

distance: for example, when the robot judges attitude at the frame t , it adopts parameters at the frame $t - s$ (few seconds before).

IV. RECOGNITION OF ENVIRONMENT IN WHEELCHAIR SUPPORT

A. Obstacle Discovery using Visual Sensors

While conveying the wheelchair, the robot must find an obstacle in the first place. The robot calculates distance of the nearest point in the real view image employing stereo measurements. If distance falls below threshold, the robot stops walking.

B. Self-Localization using Recognition of Wheelchair

In wheelchair support, it is necessary for the robot to perform self-localization. This paper deals with recognition of the wheelchair as an example of self-localization.

There are many cases that the robot separates from the wheelchair: for example, to pick up the thing on the ground the robot leaves the wheelchair for it.

When the robot return to the wheelchair according to the 3D model of the environment, the robot ca not come back to exactly the same place because of difference between the model and the real environment. Therefore, the robot must necessarily recognize the wheelchair before holding a grip bar of it.

Our work employs the model-fitting method using multiple visual cues and matches the three dimensional model of the wheelchair with the real world wheelchair. The robot utilizes not the whole wheelchair model but the bar model of the wheelchair in recognizing. The reason is follows:

- 1) Since the robot is near the wheelchair so as to push it, it is difficult to recognize the whole wheelchair.
- 2) The robot only touches the bar. Hence it is not necessary to recognize the whole wheelchair.

We introduce the particle filter technique for model-fitting [7]. The particle filter provides the robust recognition of time-varying state vectors. Okada et al. researched integration between the recognition using multiple visual cue and behavior of the humanoid robot [8]. The detail of the recognition according to [8] is follows.

To recognize the wheelchair we employ two visual cues: 3D feature points and 2D straight edges. The two visual cues above give following probability:

$$p(z_t|x_t) = p_{point}(z_t|x_t) p_{edge}(z_t|x_t) \quad (9)$$

Let z_t be the measurement vector using some visual cue at the frame t . x donates the state vector representing position and rotation of an object. $p(x|z_t)$ stands for the posterior distribution. x is usually the vector with 6 elements but in this case represents three dimensional vector(Fig.5). On recognizing the wheelchair the state vector define as $x_t = [x_t \ y_t \ \theta_t]^T$. That is the robot estimate horizontal position(x_t , y_t) and yaw rotation(θ_t) at the frame t . The measurement vector z_t is visual cues.

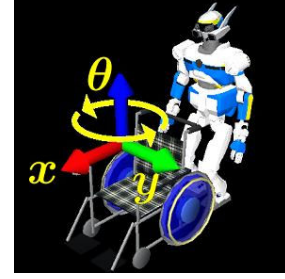


Fig. 5. Model of the system. The state vector is displayed around an arrow.

The likelihood using 3D feature points is defined as

$$p_{shape}(z|x) = \exp\left[\frac{(\frac{1}{|P|} \sum_{p \in P} D_{point}(p, F_{ref}^{visible})^2)}{2 \sigma_{shape}^2}\right]$$

$D_{point}(p_1, p_2)$ denotes the squared 3D distance between two points. σ_{shape} is a user defined weight value. Here let $F_{ref}^{visible}$ be a set of visible faces of the wheelchair model from the robot's view point. P is a set of 3D feature points near the $F_{ref}^{visible}$ and $|P|$ denotes the number of P .

The likelihood using 2D straight edge as below.

$$p_{edge}(z|x) = \exp\left[\frac{(D_{edge}(E^{2D}, E_{ref}^{2D}))^2}{2 \sigma_{edge}^2}\right]$$

The wheelchair has a lot of 2D straight edge. First we apply the Canny edge detection method to extract reliable edges from a input image, then we extract straight edges E^{2D} near the farthest point.

E_{ref}^{3D} , which is associated with the wheelchair model, are projected on the image plane to gain the 2D reference edge E_{ref}^{2D} .

$D_{edge}(E_1, E_2)$ represents measurement of the similarity between E^{2D} and E_{ref}^{2D} [8].

Fig.6 illustrates the model-fitting experiment.

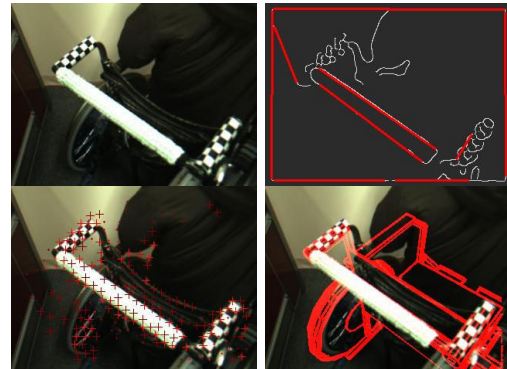


Fig. 6. The wheelchair model fitting. First : the robot's view :: Second : 2D straight edge :: Third : 3D feature points :: Fourth : the result image in which the wheelchair is located in the estimated position and rotation, superimposed with red lines.

Fig.6 shows that the robot can recognize the wheelchair.

V. HUMAN-INTERFACE BEHIND THE USER IN WHEELCHAIR SUPPORT

In many cases of nursing care the humanoid robot is located near a person. In wheelchair support the robot stands behind a person and under such condition as back-interaction:

- 1) Hard for the robot to recognize human face and exact gesture from a person.
- 2) Difficult for the person to look around and to view behavior of the robot.

In this section, we treat back-interaction as following communication:

- From the person to the robot
Face detection and gesture recognition behind the person
- From the robot to the person
Informing the person about behavior of the robot using the robot's speaker

A. Face Detection behind the Person

While pushing the wheelchair the face of the person is not usually visible to the robot. However, it is important that the robot detects when the person is looking back and recognize him/her. This looking-back motion will be the cue when the person wants to communicate with the robot. In this subsection we describe the face-detection method using particle filter. In this section, we apply particle filter to face detection. The modified point of the particle filter from Section IV-B is follows:

- 1) State vector
 $\mathbf{x}_t = [x_t \ y_t \ z_t]^T$ is 3D position of the face.
- 2) How to calculate likelihood
In calculating likelihood, we employ 3D position and color histogram of the face at the frame t .

The likelihood using 3D position is defined as follows:
If ($X_{max} > x_{t|t-1} > X_{min}$) *and* ($Z_{max} > z_{t|t-1} > Z_{min}$)

$$p_{pos}(z_t|x_t) = \exp(-\alpha(|x_{t|t-1} - x_{t-1|t-1}|))$$

else

$$p_{pos}(z_t|x_t) = 0$$

The likelihood using color cue is defined as follows:

$$p_{batt}(z_t|x_t) = \exp[-\alpha_{color}(B^2(h_{B_x}^H, h_{B_{ref}}^H) + B^2(h_{B_x}^S, h_{B_{ref}}^S) + B^2(h_{B_x}^V, h_{B_{ref}}^V))]$$

$$p(z_t|x_t) = p_{pos}(z_t|x_t)p_{batt}(z_t|x_t)$$

α is a positive constant. In order to omit particles far from the wheelchair, X_{max} , X_{min} , Z_{max} , Z_{min} is defined.

The similarity of two color histograms in Equation 10 is measured using Bhattacharyya coefficient. For measurement of Bhattacharyya coefficient, we compare HSI color histogram [9] about small regions around every points on the view image (h_{B_x}) with knowledge-based face HSI color

histogram ($h_{B_{ref}}$). $B(\mathbf{h}_1, \mathbf{h}_2) = [1 - \sum_{b=1}^{N_b} \sqrt{h_{b,1}h_{b,2}}]^{1/2}$ is Bhattacharyya distance [10].

Fig.7 is the face detection example. This images shows convergence of particles and implies that detection of the face succeeds.

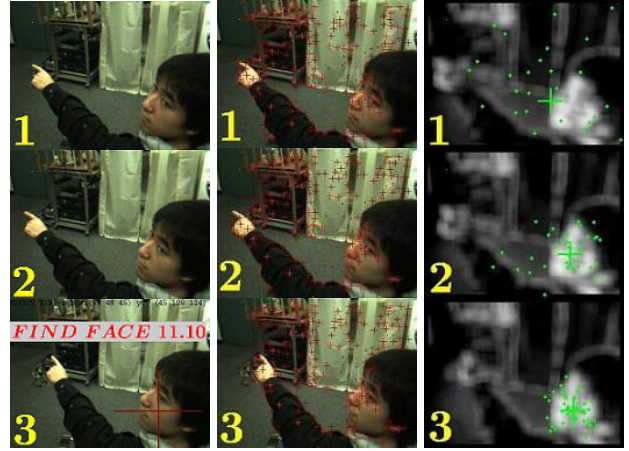


Fig. 7. Face detection using particle filter. The left images is the view image of the robot, the center images displays KLT feature points and the right images is the color-extracted view image. At the right images, the whiter pixel has higher similarity with face color histogram. Also, small green crosses express particles and the large green cross displays the estimated center of the face at the right images. Suppose summation of particles' weights exceed a threshold, "FIND FACE" is displayed.

B. Gesture Recognition behind the Person

This paper assumes behavior of the humanoid robot instructed by a human: while pushing the wheelchair, a human tells the robot where to go. The robot must recognize human gesture when the human instructs the robot not only to start and to suspend but also to move to desired directions.

In wheelchair support, fingers of the person is not easily visible from the robot. Therefore it is difficult to recognize fingers clearly.

In this section, we treat following instructions:

- 1) Emergency stop indication
- 2) Direction indication (Go Forward, Turn Left, Turn Right)

The person points desirable directions for instruction. If the person wants the robot to suspend walking, it will be sufficient to raise his/her hand.

We explain the algorithm of direction indication in follows. The robot knows gestures by estimating hand position and shoulder position from 3D feature point. Detail processes is follows:

- 1) Filter of features

S let a set of 3D feature points.

$$S' = \{f \in S \mid \text{if } f \text{ is on the wheelchair}\}$$

- 2) Estimate position of the hand

Sort S' in descending order about X -component.

Measure the center point using elements $f'_1, \dots, f'_k \in \text{sorted } S'$ and let the point be estimated hand position

p_{hand} .

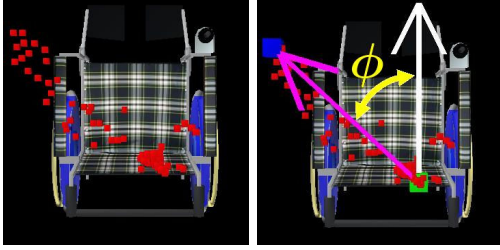


Fig. 8. Estimation of position of the hand and the shoulder(left : before estimation :: right : after estimation). At the right image the blue rectangle display p_{hand} and the green one $p_{shoulder}$.

3) Estimate position of the shoulder

Divide S' according to whether an element is $Y > p_{head}$ or $Y < p_{head}$ and extract the larger set and leg the set be S'' .

Sort S'' in descending order about Y -component If S'' is the set of which element is $Y > p_{head}$ otherwise in ascending order.

Measure the center point using elements $f_1'', \dots, f_i'' \in$ sorted S'' and let the point be estimated shoulder position $p_{shoulder}$.

Here, X -component and Y -component obey the humanoid's coordinates.

4) Calculating instructive direction

To calculate ϕ , which is instructed direction, according to p_{hand} and $p_{shoulder}$ (at the right image of Fig.8).

$$\phi [degree] = \frac{180}{\pi} \text{Arccos} \left\{ \left(\frac{p_{shoulder} - p_{hand}}{|p_{shoulder} - p_{hand}|}, e_x \right) \right\}$$

Here let (\cdot) be inner product and let e_x be X unit vector.

Next we explain the algorithm of emergency stop indication.

1) Filter of features

$$S''' = \{ f \in S \mid \text{if } f \text{ is in the robot's presence} \}$$

2) Check whether suspension indication exists or not

If S''' is not an empty set, recognize emergency stop indication.

VI. EXPERIMENT AND RESULT

This section shows the wheelchair support experiments using the above-mentioned technique.

A. The Experiment on Picking-up and Handing-on Motion

Fig.9 is the experiment that the robot recognizes the wheelchair.

- (1) Recognize the pointed direction
- (2) Pick up an object on the ground
- (3) Return to the wheelchair and recognize the wheelchair
- (4) Hand the object to the person

In (3), the robot recognizes using the above-mentioned technique(Section IV-B). This experiment shows that recognition of the wheelchair enables the robot to behave apart

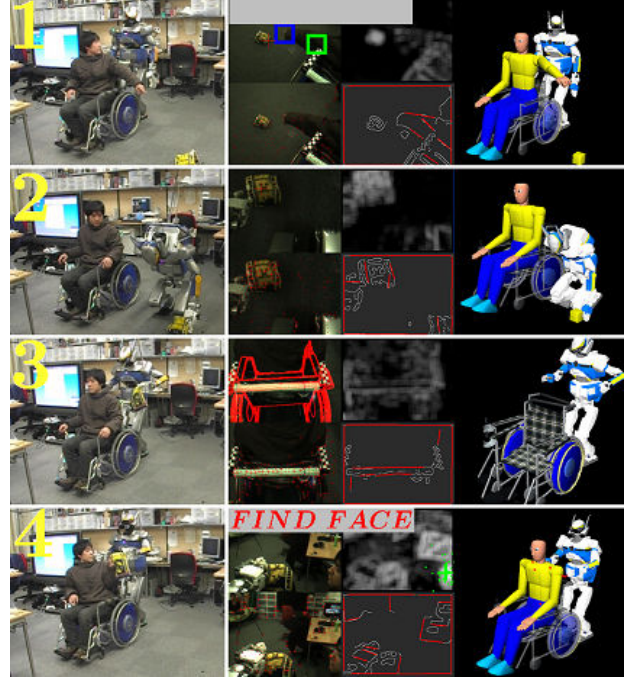


Fig. 9. Wheelchair-recognizing experiment

from the wheelchair and to resume control of the wheelchair. In this experiment, we apply the recognition of the user's looking-back motion to the detection of the face. Thus, the robot hands the object to the person after the detection of the face. This result implies that the robot can know when to hand the object to the person by using the detection of the face.

B. The Experiment on Human-led Wheelchair-Pushing

Fig.9 is the experiment that the robot detects obstacles and obeys the person's instructions.

- (1) Recognize the "Go Forward" instruction and start walking.
- (2) Detect the rough terrain and stop walking. Recognize the "Go Forward" instruction and resume walking after the suspension.
- (3) Detect the obstacle on the ground and stop walking. Recognize the "Go Backward" instruction and resume walking after the suspension.
- (4) Recognize the "Stop" instruction and stop walking.
- (5) Recognize the "Turn Right" instruction.
- (6) Turn right.

In (1), the robot starts walking on an even ground. Since the robot learns parameters of attitude distribution on an even ground, on the rough terrain the Maharanobis distance(Equation 8) will exceed the threshold(in this experiment the threshold is 2.5).

In (2), the robot resumes walking on the rough terrain and learns parameters of attitude distribution on the rough terrain. After the wheelchair collides with an obstacle in (3),

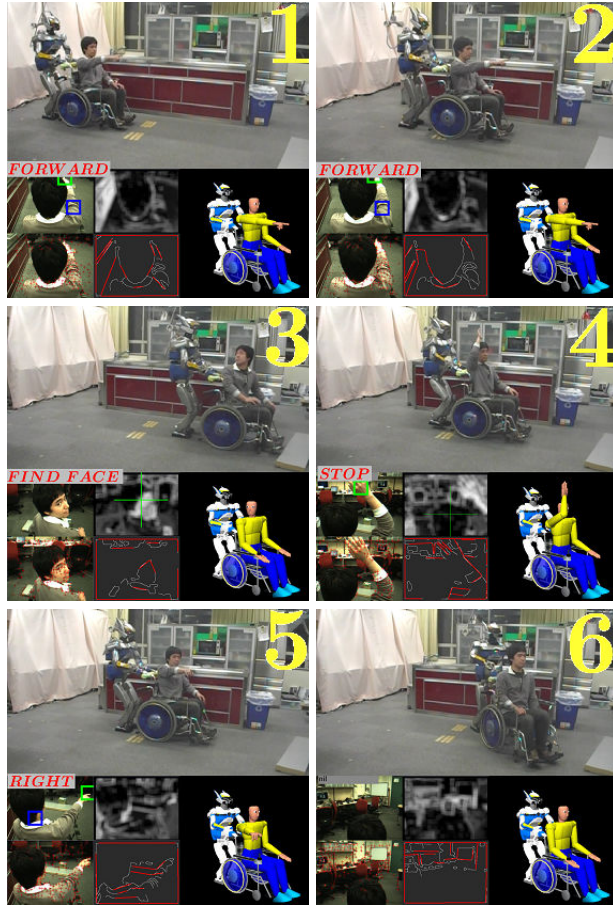


Fig. 10. The experiment on human-led wheelchair-pushing

the Maharanobis distance exceed the threshold again and the robot suspends walking.

In this experiment, the robot recognizes the following gestures: “Go Forward”, “Turn Right” and “Stop”. In this case, the “Go Backward” instruction is applied to the detection of the face. This experiment shows that integration of recognition of human instruction enables the robot to push the wheelchair to the desired direction and avoid the obstacle.

VII. CONCLUSION

This paper takes up several tasks in wheelchair support and presents following antidotes to achieve each task.

1) Whole-body motion

Compensation of the ZMP in proportion to reflected forces enables the robot to transfer the wheelchair. While conveying the wheelchair, the robot discovers such abnormality as collision with an obstacle according to the outlier-detection of the attitude.

2) Recognition of environment

The robot finds obstacles by stereo measurement. Recognition of the wheelchair using particle filter makes readily achievable humanoid’s behavior separate from the wheelchair.

3) Interaction behind a person

The robot obeys a person’s instruction by face-detection using particle filter and recognition of gesture estimated by arrangement of 3D feature points. On switching behavior the robot tells content of behavior to a person.

Since wheelchair support includes many tasks, the integration of these elemental technique is quite vital. In our work, we integrated 1, 2, 3 and so on. In this paper, we constructed foundation of integrated system for wheelchair support.

REFERENCES

- [1] Y. Sakagami, R. Watanabe, C. Aoyama, S. Matsunaga, N. Higaki, and K. Fujimura. The intelligent ASIMO: System overview and integration. In *Proceedings of the 2002 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS'02)*, pages 2478–2483, 2002.
- [2] R. Dillmann, P. Steinhaus, and R. Becher. ARMAR II - A Learning and Cooperative Multimodal Humanoid Robot. *International Journal on Humanoid Robotics*, 1(1):143–156, 2004.
- [3] M.Inaba, S. Kagami, F. Kanehiro, Y. Hoshino, and H. Inoue. A Platform for Robotics Research Based on the Remote-Brained Robot Approach. *The International Journal of Robotics Research*, 19(10):933–954, 2000.
- [4] K. Okada, T. Ogura, A. Haneda, D. Kousaka, H. Nakai, M. Inaba, and H. Inoue. Integrated System Software for HRP2 Humanoid. In *Proc. of International Conference on Robotics and Automation (ICRA'04)*, pages 3207–3212, 2004.
- [5] S.Kajita, F. Kanehiro, K. Kaneko, K. Fujiwara, K. Harada, K. Yokoi, and H. Hirukawa. Biped Walking Pattern Generation by using Preview Control of Zero-Moment Point. In *Proceedings of the 2003 IEEE International Conference on Robotics and Automation*, pages 1620–1626, 2003.
- [6] K.Yamanishi, J.Takeuchi, G.Williams, and P.Milne. Online Unsupervised Outlier Detection Using Finite Mixtures with Discounting Learning Algorithms. In *Proceeding of ACM SIGKDD Int'l. Conf. on Knowledge Discovery and Data Mining*, pages 320–324, 2000.
- [7] Michael Isard and Andrew Blake. Condensation – conditional density propagation for visual tracking. In *International Journal of Computer Vision*, pages 5–28, 1998.
- [8] Kei Okada, Mitsuharu Kojima, Satoru Tokutsu, Toshiaki Maki, Yuto Mori, and Masayuki Inaba. Multi-cue 3D Object Recognition in Knowledge-based Vision-guided Humanoid Robot System. In *Proceedings of the 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS2007)*, pages 3217–3222, 2007.
- [9] Patrick Perez, Carine Hue, Jaco Vermaak, , and Michel Gangnet. Color-based probabilistic tracking. In *ECCV '02: Proceedings of the 7th European Conference on Computer Vision-Part I*, pages 661–675, 2002.
- [10] T. Kailath. The Divergence and Bhattacharyya Distance Measures in Signal Selection. In *IEEE Transaction on Communication Technology*, pages 52–60, 1967.