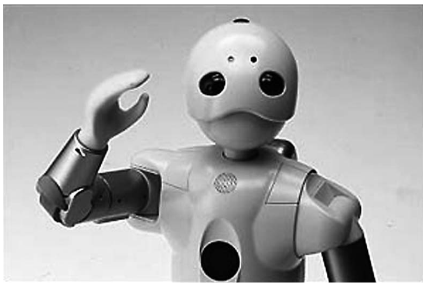


Computer Vision Technologies for Home-use Robot "wakamaru"

TETSUYA TOMONAKA*1 YOSHIHIRO KOKETSU*1
RYOHTA HIURA*2 KEN OHNISHI*2
KIICHI SUGIMOTO*3

1. Introduction

Mitsubishi Heavy Industries, Ltd. (MHI) has been developing the home-use robot "wakamaru" as a robot that lives with people(1). The concept is that of a robot having daily rhythm, that can talk, live as a family member and play a useful role in home life.

The robot "wakamaru" is equipped with various sensors as shown in **Fig. 1**. The robot locates and speaks to people, and watches the house when no-one is at home. When the battery level becomes low, the robot goes to the charging station by itself to charge itself up automatically, thus behaving independently around the clock.

In the case of a home-use robot, unlike an industrial robot, the working environment cannot be changed for the robot's convenience. In the event of disturbances such as climate changes or lighting fluctuations, the robot is expected to recognize such changes and behave appropriately and safely.

By combining the image data taken by two cameras (a human recognition camera installed in the robot's forehead and an omnidirectional camera in the top of the robot's head) and the detection data obtained by various sensors on the robot's body, constant "awareness" of its own position and detection of human figures can be realized. An outline is given in this paper.

2. Image recognition in home environment

In an ordinary household, the image obtained by the camera varies significantly due to sunlight conditions, lighting changes and the effects of people living in the house. In a living room having with a big window, strong sunlight enters and significant changes occur depending on the season, weather and time, while other disturbances are caused by the opening or closing of curtains, or turning on or off the lights.

(1) Self-localization

For the robot to move and behave independently in the home, it must determine its own position and movements. Estimation of its position by odometry (to determine the travel distance or the turning angle from the rotational angle of wheels) may be subject to cumulative errors due to slipping of wheels. In particular, a slippery carpet will increase the error range.

Therefore, the information obtained by the omnidirectional camera such as that related to the surrounding wall or ceiling is registered in advance, and the robot corrects its position by referring to the recorded information when moving.

The position must also be estimated correctly even if the sunlight or lighting conditions change significantly, or the wall or ceiling is shielded by the people in the room.

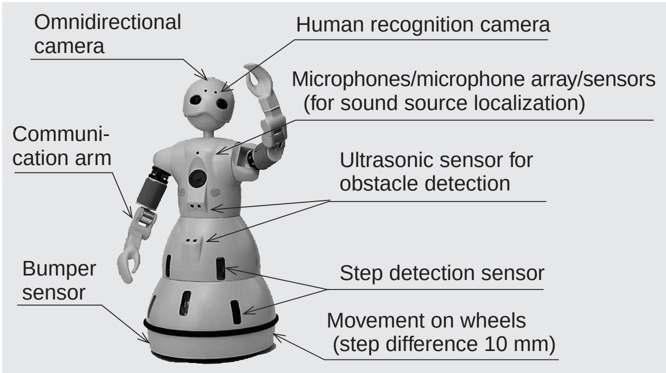


Fig. 1 Configuration of home-use robot "wakamaru"

(2) Detection and recognition of human

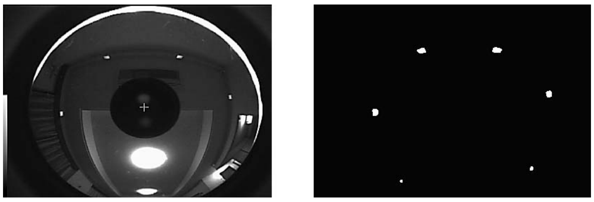
The robot "wakamaru" detects a moving object by the omnidirectional camera, moves toward it, finds the face by the human recognition camera, approaches it and recognizes the face. By this series of actions, an individual person is identified, and the robot speaks to the person on appropriate topics.

In moving object detection technology by a conventional method, a swinging curtain could be mistaken for a person. And a lamp (or a small object) resembling skin color could be mistaken as a human face. By increasing the reliability of the individual technologies and combining a series of processes, human figures can now be detected and recognized more accurately.

3. Robot status estimation by probabilistic model

A home-use robot works in an environment that cannot be changed freely, unlike the case of an industrial robot. The data from various sensors and image processing results cannot be handled individually and must be combined to enhance the reliability. The robot "wakamaru" estimates its status by employing the particle filter(2) technique (that estimates a status sequentially using a probabilistic model). This enhances the reliability of environmental recognition by combining various processing results.

*1 Advanced Technology Research Center, Technical Headquarters
*2 Kobe Shipyard & Machinery Works
*3 Takasago Research & Development Center, Technical Headquarters



(a) Input (b) Landmark extraction result

Fig. 2 Example of landmark extraction by omnidirectional camera

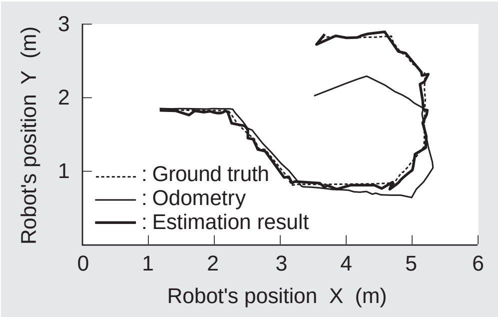


Fig. 3 Example of self-localization

In the status estimation, the problem is formulated to estimate the probability density function p (Xt | Zt) of status vector Xt. Where, the status vector Xt shows the three-dimensional position of a human (or a robot) at time t, the observation vector Zt shows the result of image processing, and the history of image processing is expressed as Zt = {Z1, Z2, ..., Zt}.

In the particle filter, the probability density is expressed by N weighted sample sets, and it is designed to estimate sequentially while updating the sample sets by the two processing steps of prediction and observation. The robot "wakamaru" executes this estimation process in real time, and realizes high reliability as shown below.

4. High reliability for self-localization

In self-localization, the position and orientation of the robot are used as status vectors and estimated by the particle filter. In the prediction step, the position of each sample is predicted by odometry. In the observation step, the detection result of landmark obtained from the omnidirectional camera (**Fig. 2**) is used as the observation vector, and the weighting is updated by determining the likelihood of each sample from the probability density of the landmark position.

Fig. 3 shows an example of the result of self-localization. By correcting the cumulative errors of odometry, a position very close to the true value is obtained.

5. High reliability for human detection

In human detection, the three-dimensional position of the human head is used as status vector, and estimated sequentially by the particle filter. In the prediction step, the position is predicted by a human motion model. In the observation step, the images by the omnidirectional camera and human recognition camera are entered as shown in **Fig. 4**, the image processing result shown in **Table 1** is used as the observation vector, and the probability density function is estimated.

(a) Omnidirectional camera (b) Human recognition camera

Fig. 4 Example of human image taken by wakamaru's cameras

Table 1 Example of image processing for human detection

Input image	Image processing
Omnidirectional camera	• Moving object detection • Human tracking
Human recognition camera	• Moving object detection • Human tracking • Skin color detection • Face recognition

Thus, in general, if there is error in individual processing or recognition, the face can be traced stably.

6. Conclusions

The robot "wakamaru" at present in the process of development, is intended to be a robot that lives with people. Household environments include various disturbance factors such as sunlight, lighting conditions, shadows cast by people, and others. It is thus important to suppress detection errors and recognition errors by combining the results of various sensors and image processing.

The robot "wakamaru" achieves human figure detection and self-localization with high reliability by estimating the status in real time using the probabilistic model, from various detection and image processing results obtained from the omnidirectional camera and human recognition camera.

References

(1) Kawauchi et al., Home-use Robot "wakamaru", Mitsubishi Juko Giho Vol.40 No.5 (2003)
(2) Isard, M. et al., Condensation-conditional density propagation for visual tracking, IJCV, Vol.29 No.1 (1998)


Tetsuya Tomonaka


Yoshihiro Koketsu


Ryohta Hiura


Kiichi Sugimoto


Ken Ohnishi