Intelligent Wheelchair Using Visual Information on Human Faces

Yoshihisa Adachi, Yoshinori Kuno, Nobutaka Shimada, and Yoshiaki Shirai

Dept. of Computer-Controlled Mechanical Systems
Osaka University
2-1, Yamadaoka, Suita-City, Osaka 565-0871, Japan

Abstract

With the increase of senior citizens, there has been a growing demand for human-friendly wheelchairs as an aid of mobility. This paper proposes a concept of an intelligent wheelchair to meet this need that can understand human intentions by observing the user's nonverbal behaviors and can move as the user wishes with minimum human operations. It also describes our experimental robotic wheelchair system. Human intentions appear most on his/her face. Thus, the experimental system observes the human face, computing the direction of the face. As the first step toward the intelligent wheelchair, we have made experiments on controlling the system's motion according to the direction of the face. Experimental results prove our approach promising.

1 Introduction

As the number of senior citizens has been increasing with the years, demand for human-friendly wheelchairs as an aid of mobility has been growing. Recently, robotic wheelchairs have been proposed to meet this need[1][2][3]. These wheelchairs help humans with the aid of ultrasonic and vision sensors to avoid obstacles, to go to pre-designated places, and to pass through narrow or crowded areas.

A wheelchair with a human on it can be considered to be a system consisting of a human and a machine. Conventional robotic wheelchairs see the outside of the system, realizing autonomous movement functions mentioned above. Although these functions are indispensable, we would like to point out the importance of seeing the inside of the system, that is, looking at the user to realize human-friendly systems. For example, if the machine can tell the user's intention of turning left in advance from his/her behaviors, it can give helpful information to him/her, such as a caution message in the case that he/she is not paying necessary attention around him/her. Or if it is certain about the user's intention, it will turn left autonomously after a simple interaction with the user for confirmation. The former method has been proposed by Pentland et al. in the case of car driving[4].

The goal of our research is to realize an intelligent wheelchair that can understand human intentions and can move as the user wishes with minimum human operations. This paper proposes our design concepts of the intelligent wheelchair and describes our experimental robotic wheelchair system. Human intentions appear most on his/her face. Thus, the experimental system observes the human face, computing the direction of the face. As the first step toward the intelligent wheelchair, we have made experiments on controlling the system's motion according to the direction of the face.

2 System design

It is a burden for humans to drive a wheelchair with much attention around them all the time. It might be helpful if the wheelchair autonomously could move toward a specified goal while avoiding obstacles. However, humans will not always go to the destination directly. They might say hello if they see their friends on the way and might drop by a cafe together. Thus a human-friendly wheelchair system should have the capability of understanding human's intention online as well as that of autonomous movements.

Figure 1 shows the configuration of our intelligent wheelchair system.

The system has ultrasonic sensors to see the outside of the system. We are planning to use the sensor data to realize autonomous capabilities such as avoiding obstacles and detecting intersections.

The system has a video camera to see the inside of the system. That is, the camera is set to look at

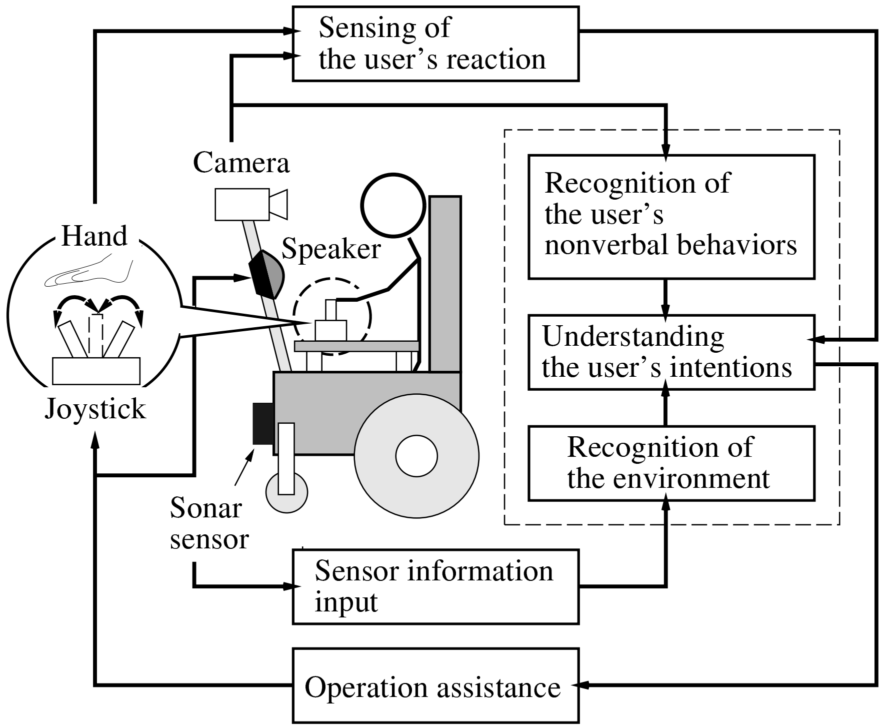

Figure 1: System configuration.

the user's face. The system understands the user's intentions from the sensor data, using them for user-friendly interface. The intention understanding cannot be perfect. Thus, we are considering the following two ways for its use.

One is to provide helpful information to the user. For example, suppose the system expects a left turn intention of the user and its ultrasonic sensing system finds an object moving fast from left behind. In this case, the system can notice the user to pay attention left-hand back by a voice message from the speaker.

The other is to move autonomously after a simple confirmation interaction with the user. The system can tell the user its intention understanding result by a message from the speaker, a flashing light, or a movement of the joystick, which has an actuator and the user can feel the force to the direction corresponding to the system's intention understanding result. The user can give confirmation by nodding or voice. In the case of the joystick, he/she can confirm by letting the joystick as the system moves it. If he/she does not want such a turn, he/she can tell the system by holding the joystick unmoved.

We are working toward the realization of the system mentioned above. We have developed a test bed by modifying a commercially available power wheelchair. Figure 2 shows our experimental system. It has a PC and a tracking vision[5] as computing resources.

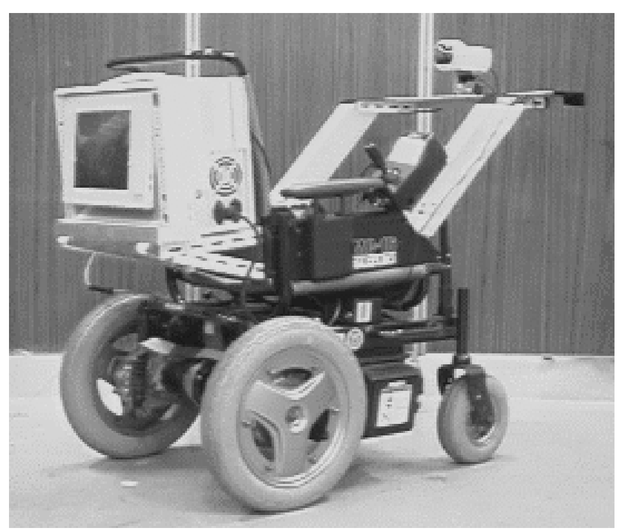

Figure 2: Experimental intelligent wheelchair system.

We consider that human intentions can be most expressed on their faces. Thus, we set up the camera to look at the user's face. We are planning to understand human intentions from human's nonverbal behaviors such as face direction (gaze direction) and facial expression. However, we cannot reach this final goal of human intention understanding by one step. Thus, as the first step toward the goal, we have developed a wheelchair system which computes the direction of the user's face and moves to the direction. For example, the system turns left if the user turns his/her head left. This is based on the assumption that the user will see in the direction where his/her wants to go. However, he/she moves his/her head in various other occasions. To avoid this problem, the current system ignores fast head movements, only responding to slow steady movements.

3 Computation of face orientation

The system calculates the orientation of the user's face, using the information to control the vehicle motion. It extracts features on the face and tracks them. Assuming that all features are on the same plane P, it calculates the orientation of the plane as the face orientation in the following way.

We use a tracking vision system[5] to track features. We choose ten feature points around the eyes, nose and mouth on the face as shown in Figure 5. It tracks these feature points based on SAD (Sum of Absolute Difference) in real time (30 frames/second).

The position data of features are sent to the host computer. It computes the motion of the plane composed of these features from the data. We assume that the motion of the human face image can be approximated by the affine motion model. We also assume that at the beginning of riding the wheelchair the user looks the camera squarely, that is, the plane P is normal to the optical axis of the camera.

Let $x_k = [x_k, y_k]^T$ and $v_k = [u_k, v_k]^T$, $(k = 1, 2, ..., 10)$ be the initial position of each feature and the motion vector of each feature (in the following explanation, we consider that the plane is fixed while the camera moves for simplicity of description). The image velocity field at a point x_k=$[x_k, y_k]^T$ can be approximated by:

$$\begin{bmatrix} u_k \\ v_k \end{bmatrix} = \begin{bmatrix} u_x & u_y \\ v_x & v_y \end{bmatrix} \begin{bmatrix} x_k \\ y_k \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}. \quad (1)$$

where the first term is a 2×2 tensor - the velocity gradient tensor - and represents the distortion of the image shape while the second term is a vector $[b_1, b_2]^T$ representing a pure translation.

We can decompose the velocity gradient tensor into three components. These components are the first-order differential invariants of the image velocity field - the curl, divergence and pure deformation components[6].

$$\begin{bmatrix} u_x & u_y \\ v_x & v_y \end{bmatrix} = \frac{curlv}{2}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} + \frac{divv}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \frac{defv}{2}\begin{bmatrix} \cos 2\mu & \sin 2\mu \\ -\sin 2\mu & \cos 2\mu \end{bmatrix} \quad (2)$$

where $curlv$,$divv$ and $defv$ represent the curl, divergence and deformation components and where μ specifies the orientation of the axis of expansion (maximum extension). These quantities are defined by:

$$\begin{aligned} curlv &= -(u_y - v_x) && (3) \\ divv &= u_x + v_y && (4) \\ (defv)\cos 2\mu &= u_x - v_y && (5) \\ (defv)\sin 2\mu &= u_y + v_x. && (6) \end{aligned}$$

We define two 2D vector quantities: A, the component of translational velocity parallel to the image plane scaled by depth Z, and F to represent the surface orientation of plane P. The orientation of the axis of expansion μ is represented with the directions of A and F[7].

$$\mu = \frac{\angle A + \angle F}{2}. \quad (7)$$

Since the plane P is normal to the optical axis at the initial position, the relation of $\angle F$ and $\angle A$ is represented as follows (see Figure 3).

$$\angle F = \angle A \pm \pi \quad (8)$$

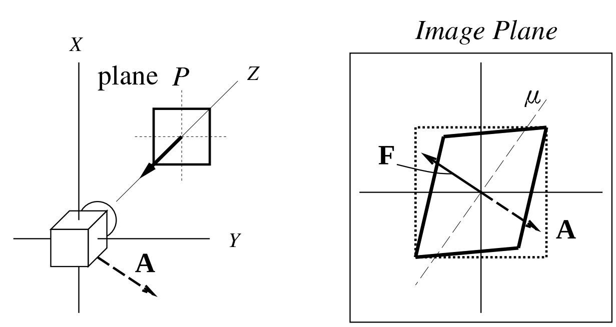

Figure 3: Camera motion.

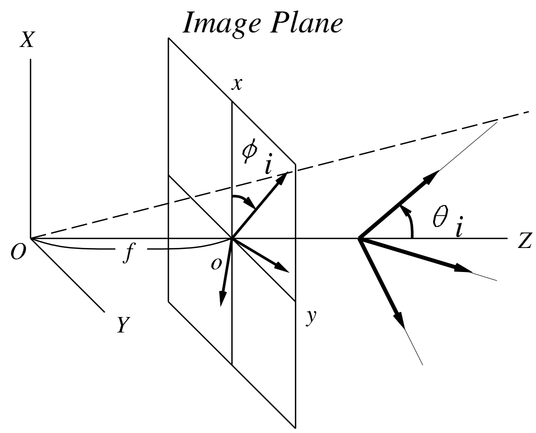

Figure 4: Unit vectors in 3D and their projections on the image.

From eqs. (5) and (6) μ is calculated and from eqs. (7) and (8) the direction of the projection of the surface orientation of plane P is calculated as follows.

$$\phi_F = \angle F = \mu \pm \frac{\pi}{2} \tag{9}$$

This equation gives two directions for $\angle F$. We solve this ambiguity as following. The camera on the wheelchair is set tilted upward by 15 degrees. Thus, if the user is sitting upright normally, the face direction seen from the camera should be a little upward. The system selects the solution satisfying this constraint if the upward tilt angle is greater than a certain threshold. If this angle is small, the user may look down a little. In such cases, the system chooses the one based on the continuity between previous direction data.

Unit vectors on plane P, $[1,0]^T$ and $[0,1]^T$ are projected onto the image by eq. (1). The angles between the projected vectors and the X-axis in the image, ϕ_1 and ϕ_2, are given by:

$$\phi_1 = \tan^{-1}\frac{v_x}{1+u_x} \tag{10}$$

$$\phi_2 = \tan^{-1}\frac{1+v_y}{u_y} \tag{11}$$

These vectors and the surface orientation of plane P are orthogonal each other in the three dimensional space. Thus, using ϕ_1, ϕ_2 and ϕ_F, the direction between the surface orientation and the optical axis is calculated as follows (see Figure 4).

$$\theta_F = \tan^{-1}\sqrt{-\frac{\cos(\phi_1-\phi_2)}{\cos(\phi_F-\phi_1)\cos(\phi_2-\phi_F)}} \tag{12}$$

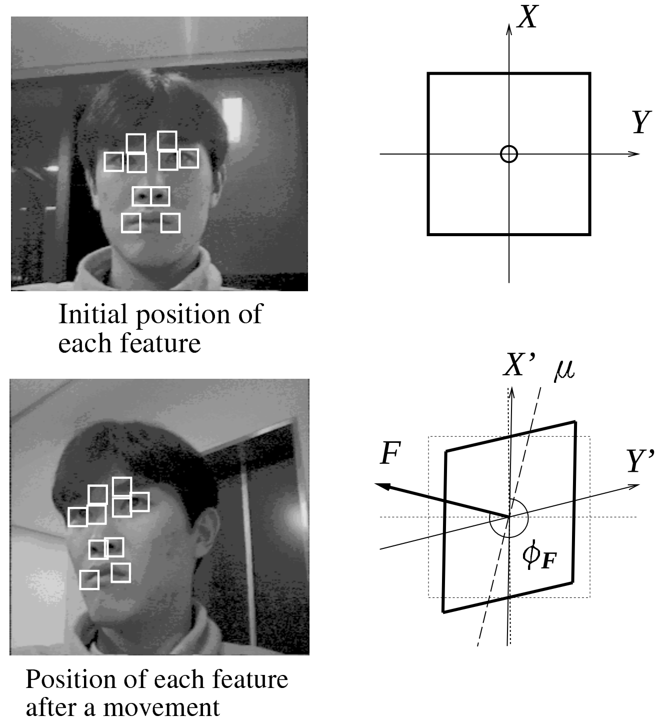

Figure 5: Estimation of face orientation.

This equation is independent of the focal length of the camera f. Thus, without the knowledge of the camera focal length, we can obtain the surface orientation of plane P, which we consider the face orientation, represented by θ_F and ϕ_F.

Figure 5 shows the position of each feature, the deformation of the image and the face orientation.

4 Experimental results

We have developed a wheelchair which moves to the orientation of the face obtained by the method described in the previous section. In this case, we only need the orientation of the face on the horizontal plane (Y-Z plane). We can obtain it by projecting the face orientation onto Y-Z plane by the following equation.

$$\psi_F = \tan^{-1}\left(\frac{\sin\theta_F\cos\phi_F}{-\sin\theta_F\sin\phi_F\sin\frac{\pi}{12}-\cos\theta_F\cos\frac{\pi}{12}}\right) \tag{13}$$

To remove small fluctuations, we apply a smoothing filter to the orientation calculation results by averaging the results of last 15 frames (data for 0.5 second). Also, we use the following rule to remove small changes of the face orientation.

If	$\bar{\psi_F} > \pi/6$,
then	turn right proportional to $\bar{\psi_F}$.
else if	$\bar{\psi_F} < -\pi/6$,
then	turn left proportional to $\bar{\psi_F}$.
else if	$\bar{\theta_F} < \pi/12$,
then	stop.
else	
then	go straight.

where $\bar{\psi_F}$ and $\bar{\theta_F}$ are the filtered results of ψ_F and θ_F, respectively.

We have made actual running experiments. Figure 6 shows an experimental scene. The left column displays images used for face orientation computation. The small white squares show the feature tracking results. The right column displays the motion of the wheelchair. The user turned the wheelchair right in the second row and left in the fourth row. Figures 7 and 8 show the changes of the face direction ψ_F and $\bar{\psi_F}$, respectively, during the experimental run displayed in Figure 6.

We made several other running experiments. These experimental results confirm that we can control the wheelchair by the face direction. Owing to filtering, the system does not respond to small and/or quick movements of the head. Thus, if we look for a while in the direction where we want to go, we can move the wheelchair to meet our intention. However, this slow response means that we cannot make quick precise control of the vehicle. We may not avoid an obstacle if it suddenly comes close. Thus, to make the system practical, we need to introduce an autonomous motion function for obstacle avoidance. The idea is that we show our intention by the face direction and after that the system can navigate autonomously while avoiding obstacles. We are now developing such a function using ultrasonic sensors.

Our final goal is to develop a human-interface system that can understand human intentions from nonverbal behaviors done by humans rather unintentionally. The current system cannot understand such intentions. It can recognize only particular nonverbal behaviors done by humans intentionally. Still, as mentioned above, we believe that the current level system is useful if we integrate the autonomous functions as well as the confirmation methods discussed in Section 2. Realizing such a system is our next step. We, then, will pursue our research to the final goal.

Figure 6: Face orientation and the wheelchair motion.

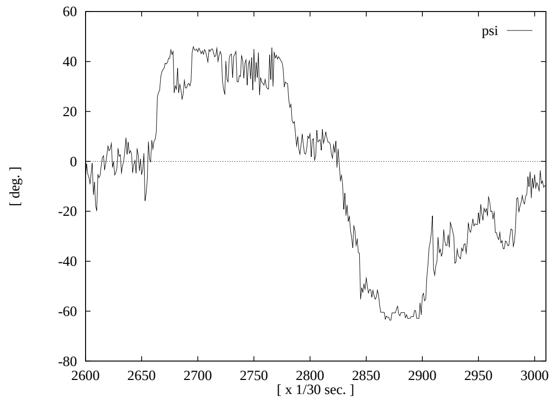

Figure 7: Face direction ψ_F in the experimental run.

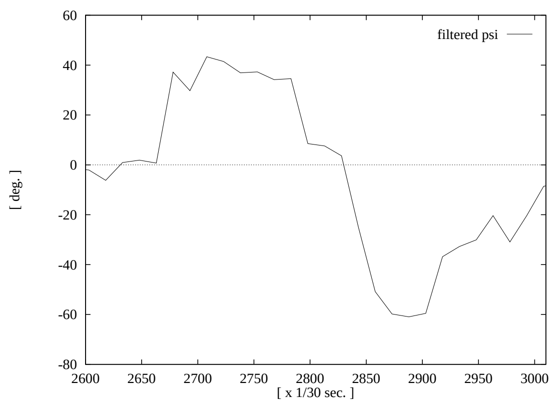

Figure 8: Filtered face direction ψ_F in the experimental run.

5 Conclusion

We have proposed a concept of an intelligent wheelchair that can understand human intentions by observing the user's nonverbal behaviors and can move by considering the intentions. As the first step toward realizing such an intelligent wheelchair, we have developed an experimental wheelchair system that can move to the direction of the rider's face computed from video camera images.

We are now working on the integration of autonomous capabilities into the current system. The sensor data for autonomous capabilities are also useful for human intention understanding. We are planning to combine them with those from observing a human to realize a more robust intention understanding method.

Acknowledgments

The authors would like to thank Hidenori Yamane for hardware design of our wheelchair system and thank Teruhisa Murashima for assisting us in our experiments. This work was supported in part by the Ministry of Education, Science, Sports and Culture under the Grand-in-Aid for Scientific Research (09555080, 09650471) and by Electro-Mechanic Technology Advancing Foundation.

References

[1] D. P. Miller and M. G. Slack: "Design and testing of a low-cost robotic wheelchair prototype", *Autonomous Robotics*, vol. 2, pp. 77-88, 1995.

[2] R. C. Simpson and S. P. Levine: "Adaptive shared control of a smart wheelchair operated by voice control", *Proc. of the 1997 IEEE/RSJ International Conference on Intelligent Robots and Systems*, vol. 2 , pp. 622-626, 1997.

[3] H. A. Yanco and J. Gips: "Preliminary investigation of a semi-autonomous robotic wheelchair directed through electrodes", *Proc. of the Rehabilitation Engineering Society of North America 1997 Annual Conference*, pp. 414-416, 1997.

[4] A. Pentland and F. Liu: "Towards augmented control systems", *Proc. of IEEE Intelligent Vehicles '95*, pp. 350-355, 1995.

[5] H. Inoue, M. Inaba, T. Mori and T. Tachikawa: "Real-time robot vision system based on correlation technology", *Proc. of International Symposium on Industrial Robots*, pp. 675-680, 1993.

[6] R. Cipolla: *Active Visual Inference of Surface Shape*, Springer, 1992.

[7] R. Cipolla, Y. Okamoto and Y. Kuno: "Robust structure from motion using motion parallax", *Proc. Fourth International Conference on Computer Vision*, pp. 374-382, 1992.