A secure and reliable next generation mobility

— An intelligent electric wheelchair with a stereo omni-directional camera system —

Yutaka Satoh * and Katsuhiko Sakaue

[Translation from *Synthesiology*, Vol.2, No.2, p.113-126 (2009)]

We propose a secure and reliable next generation smart electric wheelchair system that is equipped with a novel 3D stereovision system referred to here as a 'stereo omni-directional camera'. The novel vision system is not only intended for use with a new generation of electric wheelchairs for conventional wheelchair users, but also for use in future advanced personal mobility devices for everyone.

Keywords : Mobility, active safety, welfare apparatus, computer vision, stereo omni-directional camera system

1 Introduction

The development of technology to improve the quality of life (QOL) of the elderly and the physically disabled is a socially important issue, and the utilization of advanced science and technology is expected. Such effort is strongly demanded socially in Japan, which is a rapidly aging country.

The mobility (the ability to move around freely or the means that allow one to do so) is an indispensable element for us. Especially when a person loses the "ability to walk," it becomes very hard to move by one's own will, causing strong limitations of daily activities. This is not the issue only for the elderly and physically disabled, but also for people who can still walk normally now, but has the possibility of losing the ability by an accident or aging. Preparing the sufficient alternative for the walking ability could be the important safety net for all of us.

The electric wheelchair is the major alternative to walking ability. Recently, it has become widely spread, and it has made it easier for people with walking-difficulty to go out freely. On the other hand, as the number of electric wheelchair increases, the accidents have increased, e.g. collision with pedestrians, obstacles and falling, capsizing at level differences and stairs. Thus, the technological developments to ensure safety are of immediate demand. In the area of automobile, the technology for driving safety is actively being developed. For example, monitoring vehicle's forward space with millimeter wave raddear and stereo camera[1] and automatic braking to avoid rear-end collision[2] are technologies that have been put to practical use. Unlike an automobile that runs on the road, an electric wheelchair is used in various situations such as crowded areas and inside a room and it is necessary to deploy next-generation sensing technology for keeping safety. Therefore, we developed an intelligent electric chair equipped with "stereo omni-directional camera" which has the ability to monitor color images and 3D information simultaneously, with no blind areas, in real time. This wheelchair helps the rider to move safely indoors and outdoors and in space shared by pedestrians. It has a basic function of detecting obstacles and level differences in all directions simultaneously, and can decelerate and stop automatically when there is some danger.

In this paper, we present the detailed picture of the process from the beginning. The picture involves the strategy, scenario, and elaborations. In what follows, we shall describe the research scenario in chapter 2. In chapter 3, we describe the selection of elemental technologies necessary for constructing the prototype, and chapter 4 presents the integration and the synthesis. In chapter 5, we describe the experiments and evaluations, and in chapter 6 the summary of this paper.

2 The research scenario

This research was conducted as part of the "Development of Assistive Technology for Safety and Convenience of the Physically Disabled People (FY 2004~2006)" Special Coordination Funds for Promoting Science and Technology, Ministry of Education, Culture, Sports, Science and Technology (MEXT) led by AIST and the National Rehabilitation Center for People with Disabilities. The rank and mission of this study within this project was "to propose the potency and necessity of advanced welfare apparatuses supported by future-oriented advanced technology." We were required not to develop a product but to present the potential for improving the welfare apparatus by advanced technology. To execute this mission, we thought it would be particularly

Information Technology Research Institute, AIST Tsukuba Central 2, Umezono 1-1-1, Tsukuba 305-8568, Japan * E-mail : yu.satou@aist.go.jp

Received original manuscript December 18, 2008, Revisions received February 12, 2009, Accepted February 12, 2009

important to transmit the information to society through papers, the press release[3], and exhibitions as an integrated and synthesized system, rather than to develop individual elemental technologies, and we decided to take the following research strategy (diagram of the model is shown in Fig. 1).

(1) Evaluate and publicize the prototype by creating it as quickly as possible, through the integration and combination of the necessary elemental technologies.
(2) Seek “elemental technologies that must be created or those that must be improved,” based on our evaluation results and the external comments obtained by publicizing.
(3) Research developments and the improvements of the elemental technologies.
(4) Return to (1).

In many cases, our researches went along the path from the creation of elemental technology, through the sophisticating stage, reaching a point concerning its applications, which is an upside-down strategy to that mentioned above. This is an attempt to consecutively create the necessary *Type 1 Basic Research* from the starting point of *Type 2 Basic Research* [4] where the elements are integrated and synthesized.

In this research, we decided not to be bound by the framework of a conventional “wheelchair” but decided to engage in R&D for a new personal mobility that can be used by any of us. Development of a device for transporting a person while sharing space with pedestrians, unlike an automobile that runs on a road carrying multiple passengers, is important in allowing all people to travel with minimum energy consumption and low pollution. Also, a large market would be developed in the future by targeting wide user range. The larger market will result in higher performance, lower cost, and better infrastructure designed for the passage of new mobility. As a result, the users of current electric wheelchairs will benefit from such developments.

For the new mobility to run safely while sharing space with pedestrians both indoors and outdoors, it is important to have a technology to sense the surrounding environment quickly and accurately and a technology to accurately detect the risks from the information obtained. In this research, by installing a stereo omni-directional camera and by advanced computer vision technologies mounted on an electric wheelchair, we attempted to realize mobility with functions that prevent collision with pedestrians or obstacles and prevent falling or capsizing from level differences.

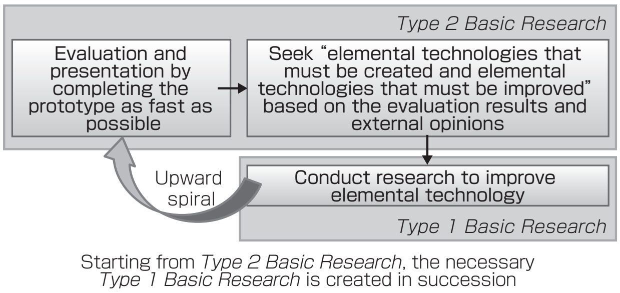


Fig. 1 Model of the research strategy.

3 Selection of elements

We shall describe the elemental technologies selected to realize the research objectives. This includes the elemental technologies that were added when we found they were necessary after the integration and synthesis of elements which is explained in chapter 4 later.

Figure 2 shows the selected elemental technologies. The function for actively sensing and recognizing the surrounding environment is realized by mounting the following onboard an electric wheelchair: (1) “sensing technology” to obtain information of the surrounding environment, (2) “cognitive technology” to detect danger in the environment from the obtained environmental information, and (3) “interface” that presents the information to the user. The individual elemental technologies are described in the following sections.

3.1 Stereo omni-directional camera

The stereo omni-directional system (SOS)[5]-[7] was developed by Satoh (one of the authors) et al., in the Human and Object Interaction Processing (HOIP) Project, Collaboration of Regional Entities for the Advancement of Technological Excellence in Gifu Prefecture, Japan Science and Technology Agency. It is an innovative camera system with a function to obtain omni-directional color image and 3D information simultaneously and in real time by assembling 36 cameras in a spherical form.

To obtain wide-range images, a camera system using a wide-angle lens and/or a parabolic mirror[8]-[10] was used conventionally. However, since it shot a wide range using just a single lens and imaging device, the optical resolution particularly of the lens was inadequate, and the resolution of the image obtained was insufficient. In contrast, the stereo omni-directional camera shoots in all directions using a large number of cameras, and high resolution can be maintained overall even if low-cost general-use lenses are used. Also, 3D information can be obtained easily by using the parallax between the cameras.

Fig. 2 Selection of elements.

Figure 3 shows the exterior view of the stereo omni-directional camera. Table 1 shows its major specifications. The collection of the individual cameras is called the camera-head, and its basic form is a regular dodecahedron. To measure the 3D information, three cameras are installed in each of the plane of the dodecahedron. Since sufficient distance is necessary between the cameras to obtain parallax (called stereo baseline, 50 mm in SOS), increasing size of the camera-head was an issue. To solve this issue, as shown in Fig. 3 right, three cameras are mounted onto a T-shaped arm (this set is called the stereo camera unit). And by arranging them three-dimensionally so each plane of the dodecahedron crosses each other without blocking the views of the cameras, the stereo baseline of 50 mm is maintained, and the camera-head can be downsized to a diameter of 116 mm or the size of a fist. The total number of cameras is 3 cameras × 12 planes = 36 cameras. All cameras are synchronized so they will shoot images at exactly the same time.

The images obtained from the camera-head are transferred to the personal computer (PC) via two optical fiber cables at 1.25 Gbps. On the PC, the 36 images are DMA-transferred to the main memory in the form where images are aligned straight, and the users are notified of the top address with a pointer. The transferred images can be accessed freely using this pointer.

By conducting a preliminary experiment of actually mounting the device on the electric wheelchair, it was found that the vibration transferred to the camera-head was greater than expected, and we strengthened the attachment of the camera-head and changed the imaging device. The initial model employed the CMOS imaging device with a rolling shutter (the shutter is released for each operation line like in a camera tube; although the structure is simple, slight distortion is produced when there is motion because of the time difference at top and bottom of the image), but slight distortion occurred in the image due to severe vibration and affected the accuracy of the 3D measurement. Therefore, we employed a CMOS imaging device with a global shutter (the shutter is released simultaneously for an entire image), since a high-performance device became available.

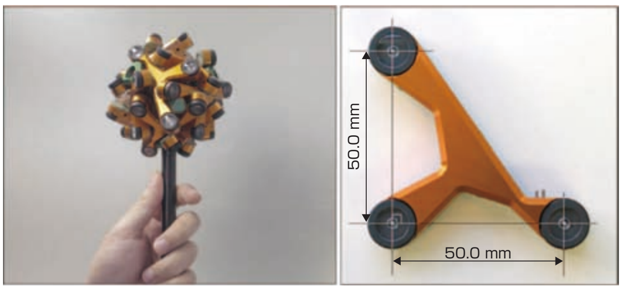


Fig. 3 Stereo omni-directional system.
The left photo shows the camerahead (diameter of 116 mm). Right photo is the stereo camera unit. The three cameras are arranged at right angles to each other in a single plane.

3.2 Stereo image processing

The distance can be calculated by the principle of triangulation from the parallax of the image shot by multiple cameras. This is like the human eyes that perceive distance using the parallax between the two eyes. Although simple in principle, there are two points that make the implementation difficult.

(1) Calibration of the stereo camera: To accurately measure the distance, it is necessary to know the actual measurements of camera parameters such as focal length, lens center, and distortion, as well as the actual measurements of positional arrangements of the multiple cameras.
(2) Search for corresponding points: Correspondence is found between points of high similarity among images shot by multiple cameras (that is, point assumed to be the same in the real world), and the distance between the corresponding points is the parallax. Objects that are near the camera have greater parallax while objects far away have smaller parallax. Since it is necessary to find correspondence for all pixels in an image, the processing cost is extremely high.

For (1), in the stereo omni-directional camera, all parameters are obtained accurately during manufacturing using the general calibration method. The camera-head has a sturdy structure so no readjustments will be necessary after manufacturing. In fact, the camera-head has been mounted on the electric wheelchair for over three years, and it has not required readjustment to the present. For (2), we considered building hardware since the processing cost was extremely high. However, considering the rapid advancement in high-speed PC, we chose implementation by software. In fact, in about three years since the commencement of the project, the computation speed of stereo image processing increased about five times purely on account of improved PC performance. Since more speed was needed with the software for it to be implemented on a small wheelchair-mountable PC, about twofold acceleration was achieved by employing parallel computation and by thoroughly removing overlapping computations.

Table 1 Major specifications of the stereo omni-directional system.

Basic form	Regular dodecahedron
Imaging device	1/4' CMOS (global shutter)
Device resolution	640 (H)×480 (V) pixels
Focal length of each camera	1.9 mm
Angle of view of each camera	101° (H)×76° (V)
Stereo baseline length	50.0 mm
Frame rate	15 fps (30 fps when color image only)
Camerahead diameter	116 mm (diameter of circumscribed circle)
Weight	About 480 g (camerahead and support)
Power consumption	About 9 W (12 V, 750 mA)

Stereo image processing can be achieved by a minimum of two cameras (binocular stereo) without any other measures such as mirrors. Yet, accuracy can be improved by using more cameras that will allow multiple results for the reliability assessment of the measurements. In the stereo omni-directional camera, trinocular stereo was used in consideration of the balance of accuracy and camera-head size.

3.3 Image integration

Figure 4 shows an example of image integration. The stereo omni-directional camera is composed of multiple cameras, and an omni-directional image is obtained by integrating the images of individual cameras using software. In our intelligent electric wheelchair, one of the functions considered for future addition is remote transmission of the omni-directional video using a cell phone line to provide remote support. Therefore, it is necessary to integrate the images in good quality, while risk detection of the surrounding environment must be done frequently, and the computation cost must be minimized as much as possible.

In general, a camera lens has greater distortion and light fall-off at the edge of the image (Fig. 5). This is not a major issue when viewing one image as in an ordinary digital camera, but gaps and differences in lightness occur at the boundaries when integrating multiple images. To solve this problem, it is necessary to: (1) correct the barrel-shaped distortion of the lens, (2) correct the peripheral light fall-off of the lens, (3) conduct geometric conversion from the coordinate system of individual cameras to the integrated coordinate system, (4) correct the color variation between the cameras, and (5) conduct the blending process for smooth connection of the boundaries between the images. Due to the limitation of space in the paper, we leave the specific computation equations to a referenced paper[11]. Since they are nonlinear conversions containing several trigonometric functions, they require over ten seconds (when 3.2 GHz CPU was used) for a single image composition. In order to improve the performance, we determined all the parameters that are dependent on the properties of the camera-head and camera unit, and performed above calculations in advance to make a transformation look-up table. With the look-up table, one finely corrected omni-directional image can be obtained from twelve raw color images with no correction at all. The time required to process an omni-directional image of 512 × 256 pixels is only 10 ms or less.

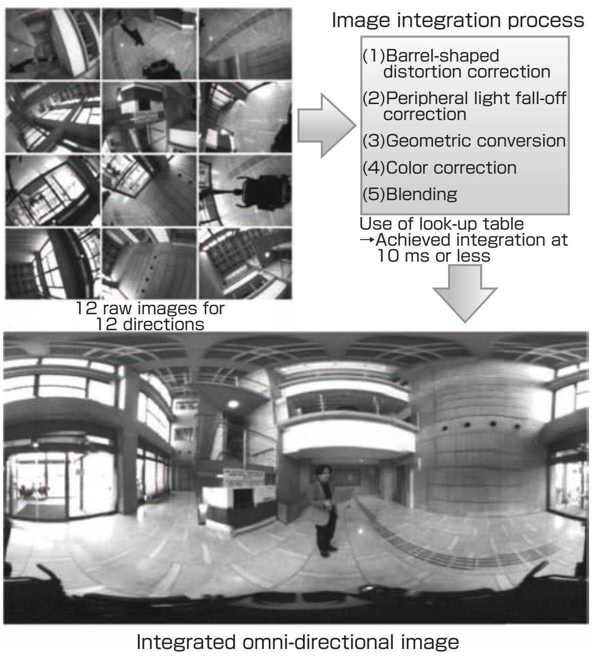


Fig. 4 High quality and high speed image integration.
The omni-directional image is produced in high quality and in high speed from 12 raw images that include lens distortion and peripheral light fall-off. Although it is difficult to present the omni-directional image in 2D, we present the image in Mercator projection, as in a world map.

3.4 Estimation of camera-head position

To obtain accurate information on the surrounding environment of the electric wheelchair, it is necessary to know accurately at what position the camera-head of the stereo omni-directional camera is attached to the electric wheelchair. In the initial design, the pose of the camera-head was obtained by detecting the direction of the gravity using an acceleration sensor fixed to the support bar of the camera-head when the wheelchair was stationary. However, two issues emerged when testing the prototype: (1) the movement of the camera-head due to unevenness and bumps during the test run was greater than expected, and it became necessary to estimate and correct the camera-head pose parameters in real time, and (2) it was discovered that a lifter (a device to hoist up the user and move him/her to the seat of the electric wheelchair) may be needed for the rider to mount and dismount the electric wheelchair. Therefore, we employed a swinging attachment arm to prevent interference of the camera-head with the lifter, but this caused slight changes of the attachment position after every swing and fix. Therefore, a method to estimate the camera-head pose parameters in real time became necessary.

Fig. 5 Lens distortion correction.
The person in the photo is holding a ruler in his hand. In the left photo, the ruler is arched since this is the image before correction. On the right is the corrected image. The distortions of the line of the ceiling and other parts are also corrected.

To estimate the camera-head pose parameters, the standard method is to use an acceleration sensor and gyroscope, but this is a relative position estimate and cumulative error is a problem. Therefore, we developed and implemented a method that allows absolute estimate of the pose of the camera-head from an omni-directional image in high speed [11]. Specifically, all edges (segments with highest brightness gradient) are extracted from the omni-directional image. Next, the directions of edges are plotted onto the voting space, and two large peaks can be obtained. This is because there are many vertical and horizontal edges in our daily environment. For example, tabletops and boundaries between the ceiling and wall have horizontal edges, while the pillars and support of the bookshelves have vertical edges. The position at which these peaks appear in the voting space presents the relative pose of the camera-head. Of course, estimation will go wrong if there is a forest of trees that grow diagonally, but the error can be detected by also using a gyroscope and an acceleration sensor.

For this method, by using look-up tables as much as possible, correction of the coordinate system and estimation of position can be accomplished in about 10 ms. Figure 6 shows an example of an image correction by actually estimating the pose of the camera-head. The upper image is an image before correction, and since the camera-head is mounted on the electric chair at a tilted position, the image seems to be distorted. The lower image is an image after conducting geometric conversion using estimated camera-head pose parameters against the same data as the one in the upper image.

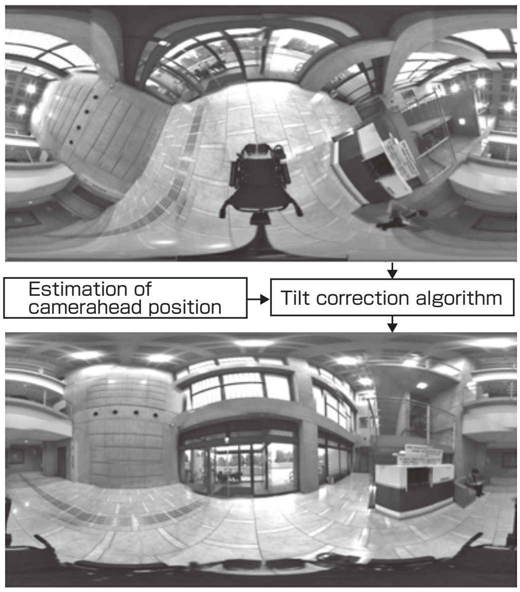


Fig. 6 Estimation of camerahead position and tilt correction.
The position of the camerahead is estimated from the vertical and horizontal edges of the omni-directional image. The upper image is the one before correction. Since it is omni-directional, the effect of the tilt appears in sine curve form. The black object in the middle is the electric wheelchair. The lower image is the image corrected according to estimated parameters. It is corrected so the sideway direction corresponds to the horizontal, and up-down direction corresponds to the vertical.

3.5 Risk detection

Figure 7 shows an example where the omni-directional distance information obtained by the stereo omni-directional camera has been visualized. The coordinate conversion is done for the distance information obtained from each stereo camera unit arranged on the regular dodecahedron, and the information is mapped onto an integrated coordinate system with the center of the camera-head as the origin. Figure 7 is an observation of the same data shot in one shot from three virtual viewpoints. Using the stereo omni-directional camera, such omni-directional distance information can be obtained 15 frames/sec (angle resolution of 360/512 degrees = approx. 0.7 degrees; about 300,000 points are shot at once). The risk detection in the environment for electric wheelchairs is conducted by directly using this omni-directional distance information. The detailed algorithm for risk detection is described in the referenced paper[11]. Basically, when the height of the floor is set at 0, all obstacles that are within the range of -0.5 m (lower than the floor) and 1.6 m high are detected. Whether the detected obstacle will be barrier to the wheelchair depends on the direction in which the electric wheelchair is moving. By setting the decision area that is switched according to the direction of the joystick, as shown in Fig. 8, when the obstacle enters the deceleration/stop area, the chair automatically decelerates or stops. In the experiment for this paper, the diameters of the decision area were 1.2 m (for deceleration) and 0.4 m (for stop). In forward straight (F_0), the decision area is rectangular to allow passage through narrow corridors. In F_{+1}~F_{+2} where the wheelchair turns while moving forward, it is expected that the amount of turn will change continuously according to the user's joystick maneuver, so a fan-shaped decision area is set to handle probabilistic spread. In the case of F_{+2} where the amount of turn becomes greater, collision at an inner radius of the direction of the turn and collision at an outer side of the turn must be considered in addition to the obstacles in the forward front direction. Therefore, the decision area is widened in the inner side of the turn, and a stop area is set in the outer side of the turn which is the opposite of

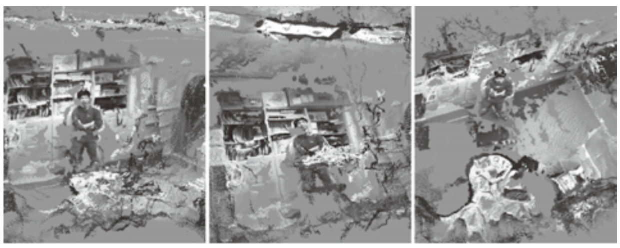


Fig. 7 Omni-directional distance information.
Same data is seen from three virtual viewpoints. There are about 300,000 observation points. Three-dimensional data for all directions can be obtained 15 times in 1 second.

the traveling direction. The observation and detection of obstacles are done in all directions at all times regardless of these decision areas, and the risk decision can be done without delay, even if the decision area is switched suddenly by sudden joystick maneuver. Moreover, since the area lower than the floor is detected, level differences and descending stairs can be detected, and the wheelchair can automatically decelerate or stop if it is judged that the situation is beyond the performance of the wheelchair. Since a bump may have rounded corners to allow passage of automobiles or carts, passage may be possible even if there is a level difference. However, if it is judged that there is a level difference of more than 5 cm, the chair is stopped unconditionally. This is because the error of distance measurement by stereo image processing may occur in the order of 2~3 cm depending on the environment condition and a decision with sufficient margin of safety is not currently possible. To solve this problem and to realize more advanced and finer risk detection, we are working separately on the development of a stereo imaging processing system using near-infrared pattern projection, and have succeeded in observing level differences of a few mm order [12]. The application of this technology to the stereo omni-directional camera will be considered in the future.

3.6 Gesture detection

We implemented the function of capturing the changes in posture and gesture of the wheelchair rider in 3D, and controlling the electric wheelchair. Specifically, (1) the function to detect abnormality of the rider's posture, and (2) the function to detect arm gesture were implemented. The space near the rider is divided into small cube regions (voxels), and recognition is conducted based on the pattern of the presence of objects within each voxels[11]. Seating position and gesturing are registered for each rider. The gesture and posture detection is done based on the comparison of the registered pattern and the observed pattern. Specific performance of the function will be described in detail in section 5.1. Although this function was not implemented at the beginning of the R&D, it was considered for implementation by request from the users after completion of the prototype.

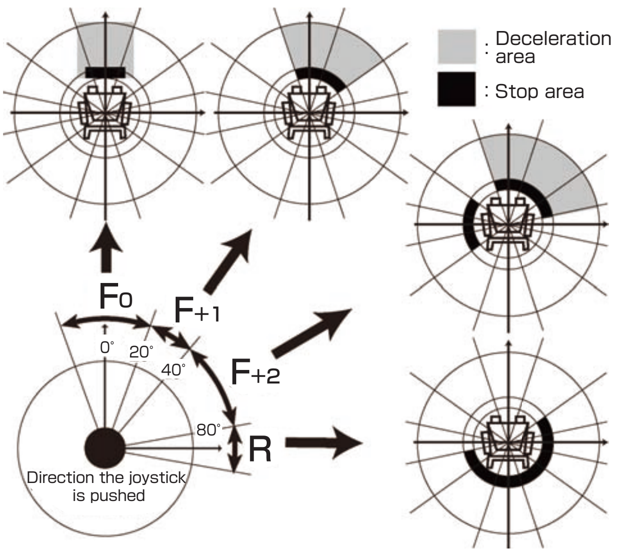


Fig. 8 Control of decision area.
The decision area is switched according to the direction the joystick is pushed. The figures take the vantage point directly above the electric wheelchair and the joystick. Top is the forward direction, and bottom is backward. Due to the limitation in space, only the forward straight (F_0) ~ right turn on spot (R) are shown, but other directions are defined in a similar manner.

3.7 Information display user interface

In an experiment using the prototype, when the electric wheelchair entered the deceleration or stop mode, the rider felt very uncomfortable when he/she could not understand the reason for the movement. Therefore, we devised a user interface to notify the state of risk detection to the rider. First, we installed a small mobile information terminal in the rider's hand region, and then considered how to display the information. In the initial stage, we thought it was better to communicate as much information as possible, and created a graphical display showing the direction and height of danger. However, it was difficult to see and understand the information that was displayed all at once when the wheelchair was moving. Therefore we devised a display that could be understood intuitively. Figure 9 left shows the display that was finally employed. The risks such as collision or fall are expressed by easy-to-understand pictograms displayed in the direction of the risk. By displaying "STOP" or "Slowdown" in large letters and in high contrast, the rider can intuitively understand "where the risk is and what kind of control is taking place."

This terminal also has a function of displaying the omni-directional image obtained by the stereo omni-directional camera. For example, it can be used to check the back view when backing up or to check the surrounding area in a bird's eye view (Fig. 9 right). The screen is a touch panel, and the desired view can be displayed by touching the screen with a finger.

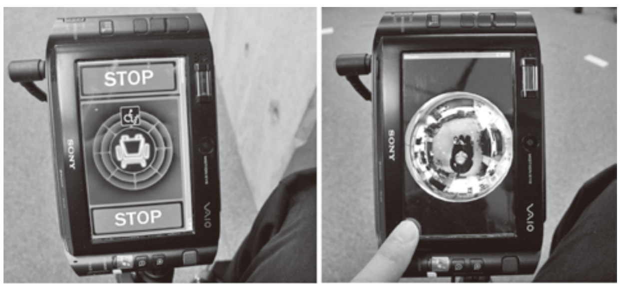


Fig. 9 Information display user interface.
The left photo shows the control status of the electric wheelchair. The type and direction of the existing risks are shown as pictograms. The right photo is the omni-directional spherical image from the vantage point of looking directly down at the rider. The sphere can be rotated freely on the touch panel.

4 Integration and synthesis of the elements

4.1 Specific process of the research

We shall describe the overall picture of the research process. Figure 10 shows the specific processes of the research based on the research strategy model diagram shown in Fig. 1. The flow is as follows. (1) As elemental technologies to realize “the sensing function” and “the danger recognition function” that are the core technologies, we initially selected the stereo omni-directional camera, stereo image processing techniques, image integration processing techniques, and risk detection techniques from conventional research ideas. (2) The integration and synthesis of the elemental technologies were conducted at the shortest time possible, the research and technology were visualized by completing the prototype, and the information was transmitted actively (through a press release and exhibitions as well as papers) to the users and to society. The exterior design of the prototype was important to increase the accuracy and efficiency of the information transmission. (3) In the evaluation phase, effort was spent on obtaining external evaluations, demands, and findings, as well as on experiments and discussions, to evaluate the current status. By doing so, the elemental technologies that must be improved or must be newly created were discerned. (4) Research for creation and improvement of the elemental technologies was conducted. This corresponds to *Type 1 Basic Research*. Since the issues that had to be solved and the evaluation standards were clear, the research could be carried out efficiently. (5) The updated or added elemental technologies were reintegrated and re-synthesized, and the cycle was repeated.

The above can be called a structure in which *Type 2 Basic Research* serves as an engine to produce *Type 1 Basic Research*, and by returning to *Type 2 Basic Research*, the research results evolve continuously in an upward spiral. Initially we expected the feedback from the user or society in the evaluation phase would be abstract and indirect (in a manner that will simply have the content or direction of the research be adjusted). However, in actual practice, the users’ evaluation became the power to directly generate the elemental technology. In the conventional research approach that originated from *Type 1 Basic Research*, users and service providers were not given much attention, but in our structure, they became the engine of *Type 2 Basic Research*, and the elemental technology generated directly became the power to propel *Type 1 Basic Research*. We feel this structure may be a new case study of “*Full Research*” [4].

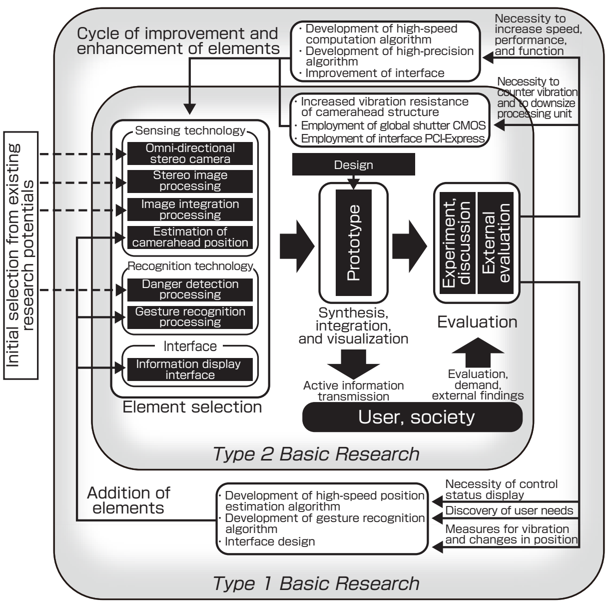


Fig. 10 Overall picture of the research

4.2 Outline of the prototype

The exterior of the prototype in which all elemental technologies described in chapter 3 are integrated and synthesized is shown in Fig. 11. This prototype has basic functions as follows: (1) the function to simultaneously detect in all directions surrounding pedestrians, obstacles, and level differences, and to automatically decelerate or stop if it determines that there is danger, and (2) the function to recognize the gesture and the posture of the rider and provide assistance.

The stereo omni-directional camera is positioned above and anterior to the rider’s head with a support arm. This position corresponds to the height of human eyes when walking, and is practical for detecting risk in the environment. Also, the position that is lower than the top height of the head is unlikely to interfere with obstacles when moving through our daily living spaces. The arm and the camera-head are kept outside the trajectory of the mounting/dismounting action of the rider. Even when lifters are used and the camera-head may be in the way, the bending part of the arm can swing away and the camera-head can be pushed to the rear.

In the initial stages of experiments with prototypes, a large

Fig. 11 Photograph of the prototype.
The stereo omni-directional system is installed in the forward position above the user’s head. All devices such as PC and power source are mounted onboard the wheelchair. It does not require external cables, and is capable of running continuously for four hours.

PC was installed outside the wheelchair and connected by cable along with an exterior power source. To mount all these devices on the chair, it was necessary to implement all functions that required large amounts of computation on just one small PC for stereo image processing, image integration, camera pose estimation and correction, risk detection, and gesture detection. To realize this, we developed and implemented various high-speed computation algorithms as described in chapter 3, and supplemented the decreased processing capacity due to downsizing the PC from the software side. We conducted thorough review of the implementations such as eliminating overlapping computation, using look-up tables as much as possible, along with introduction of parallel processing. In hardware, to deal with a small motherboard, we changed the interface of the stereo omni-directional camera from PCI-X to PCI-Express standard. Although the PCI-X standard has wide bandwidth for data transfer, it is limited for use in a server motherboard. PCI-Express standard is recently being used widely and can be used in almost all motherboards, and can maintain sufficient bandwidth for data transfer of the stereo omni-directional camera.

Figure 12 shows the exterior view of the PC mounted on a wheelchair. The PC fits compactly in the cover behind the seat. The interface unit for the electric wheelchair and PC, the joystick interface unit, and the wireless LAN device are all fitted under the seat. Lead battery (12 V, 52 Ah × 2) for driving the motor is used as the power supply for other devices as well. This requires no exterior cables, and about four hour continuous operation is possible.

Fig. 12 Mounted PC.
Right photo shows the device exposed without the cover. The PC is mounted behind the seat.

4.3 Exterior design

To create something that will "present the potential and necessity of advanced welfare device through future-oriented advanced technology" according to the research scenario, the product must draw people's interest and convey the concept and future potential. It is necessary to be concerned about the exterior design, and our team repeated discussions and created prototypes. Figure 13 left shows a scene during prototype creation. The right photo is a design for Prototype 1, and a press release[3], exhibitions, and demonstration on the TV news were done using this design (currently, the devices mounted on the wheelchair are downsized further; Prototype 2 shown in Fig. 11 is also in operation).

The points of the design are as follows. (1) The mechanics and cables should be covered as much as possible. To show them exposed will express that the product is "still at the level of laboratory experiment" and that will not be appropriate for the mission. (2) The cover should not be forcedly made small with no consideration for the design theory. If the cover is made small to appeal that the mounted devices are small, the overall balance will be lost and the covered part will stand out as an "extra area." For example

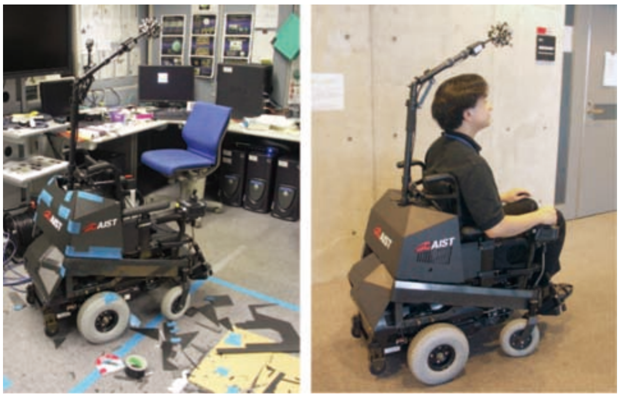
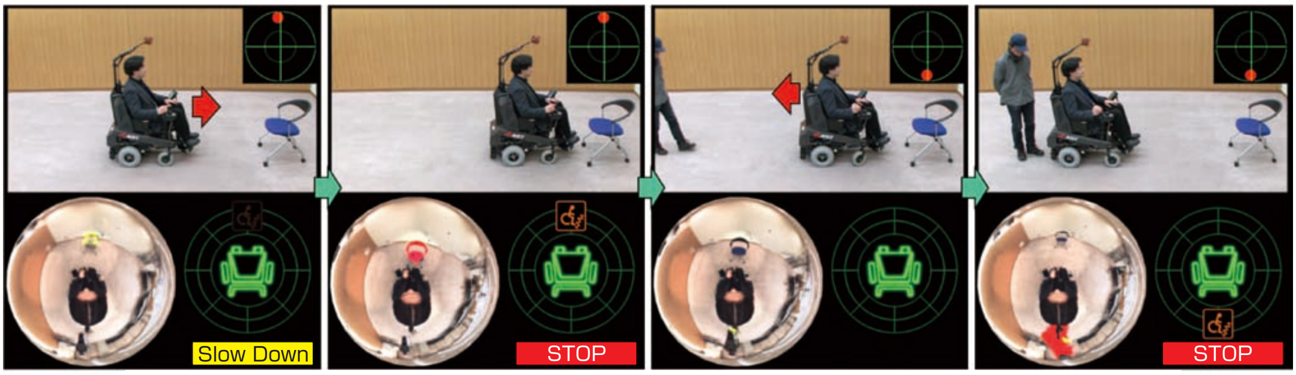


Fig. 13 Consideration of the exterior design.
Left photo shows the process of designing. The design was determined after repeated trials. The right photo is the design for Prototype 1. This model was used for the press release.



(1) (2) (3) (4)

Fig. 14 Detection of obstacles.

in Fig. 13 right, the volume filled by the mounted device is about one-third of the volume of the cover. However, it does not stand out since there is an overall good balance. We even received a question, “Where is the computer?” (3) To show that this project is an AIST effort, the AIST logo is included in the design.

5 Evaluations

5.1 Experiments using the prototype

Figure 14 shows the basic obstacle detection experiment. (1)~(4) show four scenes in chronological order. For each scene, the direction of the joystick operated by the rider is shown in the right-top (top direction is the forward direction), the omni-directional image shot by the stereo omni-directional camera in the left-bottom (expressed as a sphere; the rider is in the center), and the screen of information display interface described in section 3.7 is shown in the right-bottom. In (1), the wheelchair approaches the obstacle (chair) in front and goes automatically into deceleration mode. Since the rider continues to push the joystick forward, the wheelchair stops automatically (2) right before collision with the obstacle. In (3), the rider pulls the joystick to start backing, but a pedestrian approaches from behind out of the view of the rider, and the wheelchair stops automatically at (4) due to danger of collision.

Figure 15 is an example of an automatic stop after detecting a staircase. Since the wheelchair detects level differences and descending stairs as well as obstacles on the street, it can prevent falls in advance.

Figure 16 and 17 show examples of gesture and posture detection function. In Fig. 16, the wheelchair stops in emergency since it detects that the posture of the rider differs greatly from the preliminary registered posture. When this situation lasts longer than preset time, it is possible for the system to automatically call for assistance by the cell phone. Figure 17 is an example where gesture detection and risk detection functions are used at the same time. When grabbing something or approaching something such as in order to press the elevator button from the electric wheelchair, if the rider cannot reach the target, the gesture of extending the arm can be used as a trigger to advance the wheelchair automatically toward the target while checking the safety. Specifically, in (1), the rider extends his arm for 3 sec. or more to grab a PET bottle, assistance begins in (2), and the electric wheelchair advances slowly. It stops automatically when the arm is retracted or before colliding with an obstacle (table in this example), and in (3) the rider succeeds in grabbing the PET bottle.

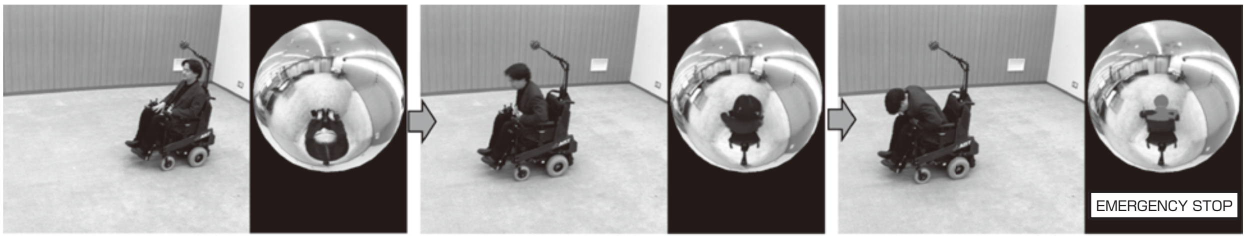
Fig. 15 Detection of descending stairs.
The descending stairs and bumps/dips are detected. The wheelchair automatically decelerates or stops if it is determined that the situation is dangerous.

This gesture recognition function determines the posture and gesture by simple comparison of the preliminary registered pattern with the 3D shape pattern roughly quantized by voxels as explained in section 3.6. This can be used only in detection of relatively large movements as in the example of Figs. 16 and 17. Some users who have disability of the arms have requested, “Can slight movement of the shoulder be recognized as a gesture?” In the future, we shall consider accurate recognition of fine movements by introducing machine learning approach.

The function shown in Fig. 17 was requested by an actual wheelchair user, and was investigated for realization. Much experience is required to fine-position the electric wheelchair using the joystick. Particularly, approaching a table to grab something or approaching a wall to press the elevator button carry large risks because of the possibility of collision with the table or wall. To avoid such problems, human assistance may be sought when fine positioning is necessary. However, such assistance is required dozens or several hundred times a day, and this makes the rider of the electric wheelchair reserved and may prevent him/her from going outdoors. Can machines support seemingly minor but high-frequency assistance? That was the users’ request and we conducted investigations.

These are the basic functions of the prototype, and we implemented several functions as their extensions. In Fig. 18, the wheelchair recognizes the nearest person, and automatically tracks the person face to face at 1 m distance. Since all directions are constantly monitored, it does not lose track even if the person makes sudden movements. In the future, we are considering a function that automatically tracks a certain person (such as an assistant) by using facial recognition technology. Figure 19 is an experiment of automatic route selection in a crowd. In Fig. 19(2), the wheelchair is surrounded by several people, but since it is observing all directions at once, it instantly decides the direction that it can take and escapes automatically in (4). In a crowd, the environment changes dynamically and constantly, and the situation may change if time is taken to gather information. Since the stereo omni-directional camera gathers information simultaneously in all directions, control is possible using the latest information at all times for all directions even in crowd situations. Here, these functions are introduced since they readily demonstrate the sensing capability of the prototype, and unlike other functions that “assist the user,” the electric wheelchair itself moves automatically. Therefore, solid safety is essential and the hurdle to practical use is high. However, the demand for such automation technology is high, and we are engaging in research for functions that must be developed for the future.

Figure 20 shows the test runs in varying environments. Test runs were conducted in various environments such as indoor space with complexly arranged furniture and outdoor space hit directly by sunlight, and stability was evaluated. The balance of “safe” and “free” became major issues in the course of the experiment. In an extreme case, the stop mode (that is, the wheelchair does not move) regardless of the user’s operation of the joystick is the “safest.” In contrast, reduced intervention by the safety system will allow “freer” movement although danger may increase. Appropriate setting differs according to the skill and physical condition of the user, as well as to simple preferences. To continue discussion of this matter in the future, we believe it is necessary to build a framework to evaluate safety in a quantitative manner, including the issue of automation mentioned before. It is also necessary to thoroughly discuss the issue of liability in case accidents do occur. The same issue exists in driving assistance for automobiles, an automatic parking function, and assistant robots. It is necessary to build social consensus



(1) (2) (3)

Fig. 16 Emergency stop when an abnormality in the rider’s seating position is detected.

(1) (2) (3)

Fig. 17 Automatic assistance is provided to the position that allows the hand to reach the bottle, by recognizing the rider’s gesture.


(1) (2) (3) (4)

Fig. 18 Automatic tracking experiment.


(1) (2) (3) (4)

Fig. 19 Automatic route selection in crowd.

by working on the framework of safety and liability while collaborating with other fields.

5.2 User response

In this research scenario, gathering the opinions of the users and incorporating them to the research was important for the upward spiral. Therefore, we exhibited in several exhibitions such as the International Home Care and Rehabilitation Exhibition (total visitors of about 100,000 people in 3 days) to hear comments of the current wheelchair users. Initially, we were concerned that “this was a rather future-oriented proposal and may not capture the interest of the current wheelchair users.” Actually, we obtained comments such as: “I have been waiting for this, I’d like one right now,” or “Please continue this kind of research.” Even when an exhibition was held in Tokyo, there were wheelchair users who came all the way from Osaka. There was a strong demand for a support system using advanced technology. Comments were provided by the visitors, and new research topics were born. As explained earlier, the automatic assistance function in Fig. 17 was produced through this process.

Since the user of welfare apparatuses is limited and customization is frequently required according to the user’s condition, it is difficult to establish as business compared to mass-produced apparatuses. Therefore, advanced technology has not been introduced despite pressing demands. However, IT and robot technologies should be actively used to meet this kind of demand, and it is necessary to consider some kind of scheme to improve the current situation. In this sense, we received lots of support and expectations from current wheelchair users for the concept of developing a new mobility rather than limiting it to conventional wheelchairs. Also, parts of the elemental technology developed in this study can be applied to safety technology in automobile industry that has a large market, and when high performance and low cost are achieved by diffusion in such a market, they can be returned and used in electric wheelchairs.

Fig. 20 Test runs in varying environments.
Test runs were conducted in varying environments, both indoors and outdoors. The left photo shows passage through a narrow corridor. Since there is almost no extra room for the width of the wheelchair, the chair stops automatically frequently when the level of safety is set high. The balance of “safety” and “freedom” is difficult. The right photo shows a test run under direct sunlight. Stability evaluations were done for various environmental changes.

6 Summary and future prospects

We described the development of an intelligent electric wheelchair with stereo omni-directional camera. As mentioned in the beginning, research and technology were visualized by integrating and synthesizing the initially selected elemental technologies based on conventional research potentials, and then by completing a “moving” prototype as quickly as possible. From the results of the evaluation and presentation of the prototype, “elemental technologies that must be created or improved” were discerned, elemental technology research was conducted, and the result was reintegrated and re-synthesized, to practice research strategy with an upward spiral structure. Necessary elemental technologies were produced consecutively. Since they were all immediately necessary and the evaluation standard of performance was clear (whether it could solve the immediately occurring problems), we were able to engage in research efficiently and with good balance.

In this research, technology for accurately and quickly sensing the surrounding environment, and the technology to accurately detect risks from the obtained information were developed to realize mobility that allows safe movement and sharing of space with pedestrians. The technologies were implemented in an electric wheelchair, and demonstration experiments were conducted. For further developments as a new mobility for all people in the future, we shall continue our investigations on many issues such as infrastructure and laws and regulations.

References

[1] H. H. Meinel: Commercial applications of millimeterwaves: History, present status and future trends, *IEEE Trans. Microwave Theory and Techniques*, 43(7), 1639-53 (1995).

[2] N. Shimomura, S. Nakamura, T. Goto, K. Fujimoto and H. Muro: A method of tracking a forward vehicle using a scanning laser radar and a camera, *Denki Gakkai Ronbunshi C (IEEJ Transactions on Electronics, Information and Systems C)*, 123(8), 1427-1438 (2003).

[3] Zen hoko sutereo kamera o tosai shita interijento dendo kuruma isu – anshin/anzen na dendo kuruma isu no jitsugen (Intelligent electric wheelchair with stereo omni-directional camera – Realization of a safe and secure electric wheelchair), *AIST Press Release* (2006) (in Japanese).

[4] *Dai 2-shu Kiso Kenkyu O Jiku Toshita Honkaku Kenkyu No Tenkai (Development of Full Research Centering on Type 2 Basic Research)*, Vol. 1~13, AIST (2006-2008) (in Japanese).

[5] K. Yamamoto, H. Tanahashi, S. Kuwashima and Y. Niwa: Sensing of real environment by stereo omnidirectional system (SOS), *Denki Gakkai Ronbunshi C (IEEJ Transactions on Electronics, Information and Systems C)*, 121(5), 876-881 (2001).

[6] Y. Satoh, K. Yamamoto, S. Kuwashima, H. Tanahashi, C. Wang and Y. Niwa: Idotai Bijon o shiko shita kogata zen hoko sutereo shisutemu (mini SOS) (Development of small stereo omnidirectional system (mini SOS) for mobile vision), *Gazo Sensing Symposium Koen Ronbunshu*

(SSII03) (Collection of Presentations of the Symposium on Sensing via Image Information (SSII03)), 311-316 (2003) (in Japanese).
[7] S. Shimizu, K. Yamamoto, C. Wang, Y. Satoh, H. Tanahashi and Y. Niwa: Moving obstacle detection by mobile Stereo Omni-directional System (SOS) using spherical depth image, *Pattern Analysis & Applications* (2005).
[8] Y. Yagi, S. Kawato and S. Tsuji: Real-time omnidirectional image sensor (COPIS) for vision-guided navigation, *IEEE Trans. Robotics and Automation*, 10(1), 11-22 (1994).
[9] J. Kurata, K. T. V. Grattan and H. Uchiyama: Navigation system for a mobile robot with a visual sensor using a fish-eye lens, *Review of Scientific Instruments*, 69, Issue 2, 585-590 (1998).
[10] C. Mandel, K. Huebner and T. Vierhuff: Towards an autonomous wheelchair: Cognitive aspects in service robotics, *Proc. Towards Autonomous Robotic Systems (TAROS2005)*, 165-172 (2005).
[11] Y. Satoh and K. Sakaue: An omni-directional stereo vision-based smart wheelchair, *EURASIP Journal on Image and Video Processing*, 2007, Article ID 87646, 11 (2007).
[12] Y. Satoh, T. Ozaki and K. Sakaue: Development of an Intelligent Wheelchair based on Stereo-vision Camera coupled with a Near-infrared Pattern Projector, *Shin Gaku Giho (PRMU2008-106) (IEICE Technical Report (PRMU2008-106))*, 108(263), 103-106 (2008) (in Japanese).

Authors

Yutaka Satoh
Completed the doctorate course at the Graduate School of Engineering, Hokkaido University in 2001. Worked as senior specialist researcher of the Human and Object Interaction Processing (HOIP) Project, Softopia Japan, and developed the stereo omni-directional camera. Currently researcher at the Information Technology Research Institute, AIST, where he engages in research for applying the stereo omni-directional camera to support physically disabled people. Also engages in research on robust pattern matching, and has done product realization of automatic human detection using a surveillance camera. Associate professor of Graduate School, University of Tsukuba (Cooperation Program between AIST and Graduate Schools). Doctor (Engineering). In this paper, worked on research planning, system design and implementation, and experiment/discussion.

Katsuhiko Sakaue
Completed the doctorate course at the School of Engineering, the University of Tokyo in 1981. Joined the Electrotechnical Laboratory in 1981, and has steadily studied image processing and its real world application. Participated in R&D for the real world intelligence technology in the RWC (Real World Computing) Project, and developed support technology for safety and comfort of the physically disabled people. Currently, senior researcher of the Information Technology Research Institute, AIST. Professor of Graduate School of Systems and Information Engineering, University of Tsukuba (Cooperation Program between AIST and Graduate Schools). Winner of Academic Encouragement Award, the Institute of Electronics, Information and Communication Engineers (IEICE) in 1979; Best Paper Award, Information Processing Society of Japan; and Fellow of International Association of Pattern Recognition (IAPR) in 2006. Doctor of Engineering. In this paper, worked on the research strategy and overall direction.

Discussion with Reviewers

1 Goal setting and scenario leading up to the goal

Question and Comment (Motoyuki Akamatsu, Institute for Human Science and Biomedical Engineering, AIST)

It is written in chapter 2 that this R&D is not limited to electric wheelchairs, but the goal is to develop new personal mobility, as you mentioned in the title. However, the scenario and technologies described after chapter 3 talk only from the perspective of application to wheelchairs. Of the functions obtained by integrating elements, please state clearly which were the functions you sought to achieve for this new personal mobility.

Question and Comment (Koh Naito, Center for Service Research, AIST)

The technologies described in the paper can be applied to things other than electric wheelchairs, and this point has been expressed in chapter 2 as well as in the title. On the other hand, the paper describes the development of various elemental technologies, their integration process, and evaluation results as *Type 1 Basic Research*, centering on the development of the prototype electric wheelchair. Please add your ideas of personal mobility society that may be constructed as this research result diffuse further, the scenario leading up to such achievements, and the technological issues that must be overcome.

Answer (Yutaka Satoh)

As you mentioned, it was unclear which of the functions among those obtained by integration of the elements were aimed at a new personal mobility. Therefore, we added in chapters 1 and 2 that the relevant points are: we are assuming the level of mobility for electric wheelchair running amidst pedestrians in indoor and outdoor spaces; and to achieve such a mobility, the "functions to quickly and accurately sense the surrounding environment and to appropriately detect the risk from the information obtained" are necessary to prevent collision with pedestrians or obstacles or to prevent fall at level differences and stairs.

Since the research was basically conducted to increase the performance of an electric wheelchair, "improvement of QOL for the elderly and the physically disabled" was an important theme from which we could not stray, and the flow of the paper emphasizes this theme. On the other hand, viewed in simple technical terms, "an electric wheelchair is a chair attached to a motor-driven cart that can be ridden by a person," we can reset the concept of a conventional electric wheelchair, and consider it as a form of new personal mobility (therefore, the exterior was designed to look different from traditional wheelchairs). We hope this may result in attracting a wider range of users, and may solve the problem of the market for electric wheelchairs. Perhaps there was a problem in mixing the QOL element and the new personal mobility element in the same text flow.

This idea has received support from the wheelchair users and the researchers at the National Rehabilitation Center for People with Disabilities, who are the counterparts of this research, and we added some comments in section 5.2. On the other hand, for the realization of new mobility, it is insufficient simply to introduce safety technology to vehicles such as electric wheelchairs, and there are mountains of issues that must be solved such as infrastructure and rules and regulations. This point was added to chapter 6.

2 Diagram on research topic setting

Question and Comment (Motoyuki Akamatsu)

You take the approach of *Type 2 Basic Research*, where you take the upward spiral research strategy, and engage in the process of prototype creation and evaluation as you set and integrate the elemental technologies. Although the specific contents are well described in the text, please show the process of setting the research topics in a diagram, so readers can understand your research process. For example, which elemental technologies did you think were necessary in the initial phase, and what problems could be solved by these elemental technologies? Also, what were the technological issues clarified through the processes of manufacturing and evaluating the prototype, and what were the perspectives of those issues (for example, durability, lack of accuracy, newly found user demand, actual use in real environment, etc.)? Please consider using something such as a block diagram to show this. Also, in the summary, please add a discussion (good points and points that must be improved) for taking the upward spiral method.

Question and Comment (Koh Naito)

The research style is described as a continuous evolution of research results by making the results visible through prototype development and efficient introduction of external knowledge. You demonstrated yourself that the various basic scientific researches (*Type 1 Basic Research*) were driven from the researches for application (*Type 2 Basic Research*). This is an important find in terms of research management, and in chapter 3, you can clarify the flow by structuring the description of each elemental technology to show what kind of *Type 1 Basic Research* was driven based on what findings. Please reconstruct the descriptions and organize them as a table at the end of chapter 3.

Answer (Yutaka Satoh)

We added Fig. 10 that summarizes the specific processes of the research. Also, we added the explanation of this diagram in section 4.1. While the specific explanations are as in section 4.1, "the cycle to enhance values" is realized when *Type 2 Basic Research* becomes an engine to actively produce *Type 1 Basic Research*, and *Type 1 Basic Research* creates or increases the performance or value of the elemental technology, and higher level *Type 2 Basic Research* is conducted using the newly created or enhanced elemental technologies. Moreover, the elemental technologies studied in *Type 1 Basic Research* are immediately needed and the result desired is clear, and therefore the evaluation standard of performance (whether it can solve the immediate problem) is clear. I think the balance and efficiency of the research was very good. We added a description on this point in chapter 6.

Initially, we thought the feedback from the evaluation phase would be indirect and abstract, and the content and direction of the research will only be slightly adjusted. Actually, the users' evaluations directly generated the elemental technologies. This was a new finding for us, and we described that process in Fig. 10. The engine for *Type 2 Basic Research* included the users and services that were ordinarily ignored in the research approach starting from *Type 1 Basic Research*. We believe that the structure in which the power to directly generate elemental technologies that will then propagate *Type 1 Basic Research* may be one form of *Full Research*.

3 Realized functions

Question and Comment (Motoyuki Akamatsu)

You introduced five main realized functions: obstacle detection, downhill slope detection, abnormal position detection, gesture detection, and automatic tracking and automatic route selection. However, you didn't provide a clear description about the situation when these functions could be useful. I think you should provide an explanation of goal setting for R&D. Are you assuming that the user may not be paying attention to his/her course of travel, are these functions set to match the level of the handicap, or are they based on the analysis of wheelchair accidents?

Answer (Yutaka Satoh)

As you indicated, the objectives and reasons for selecting each function were not clear, so we added descriptions in chapters 1, 2, beginning of 3, Fig. 2, and section 4.2. The assumed users are all people including non-physically disabled people. In considering safe runs in an environment shared with pedestrians, we thought "accurate and quick sensing of the surrounding environment and accurate detection of the risk from information obtained" were mandatory, and implemented obstacle and level difference detection as a priority item.

When we gathered information from the elderly and the physically disabled, we found there was a demand for gesture recognition, and implemented this function in the process of an upward spiral cycle (this is described in section 3.6). The functions of automatic tracking and automatic route selection were studied as additional efforts for future mobility. They were described in the paper since they demonstrated the sensing capacity of the prototype in a readily understandable manner. Yet their positioning was unclear, so we added an explanation in section 5.1.

4 Relationship to other automation technologies

Question and Comment (Motoyuki Akamatsu)

At the end of section 5.1, there is a discussion on safety and freedom. About 10 years ago in the field of automobile ITS, much were discussed about test runs on actual roads for automatic driving. In case of automobiles, the potential for damage was high, and the point of discussion was who would be responsible in case an accident occurred. After all, if an accident occurred under complete automatic drive, the manufacturer would likely be held liable, so the direction shifted to driving support technology where the driver will be in control. The strategy is to maintain a situation where the driver is always involved, and the technology will be introduced to society as an assistance technology for actions initiated by the driver. In the process of diffusion of assistance technology, the reliability of the system and the performance may increase, and society may become more willing to accept the automation system. As a driving support system centering on sensing technology, the collision warning system is starting to diffuse into the market. Here, we are discussing the design of user interface (how to communicate the detected situation to the user accurately and quickly) and the overconfidence in a warning system (such as distraction because the user depends on the warning system). In the ITS field, the problem of safety and freedom are discussed in terms of role division of the system and the user for control, and overconfidence and user interface design for sensing.

This R&D is based on sensing technology, and I think the point of introduction to society is how to balance the control by user him/herself and the control by the system, as mentioned above. Please include your comment on this point.

Answer (Yutaka Satoh)

This is a complicated issue, but I think the case of ITS which you described presents the situation accurately. We shall aim for a society where automation is accepted, by increasing the reliability and performance of the system. Although we do not have a specific plan yet, we believe it is necessary to actively build a framework for evaluating safety in collaboration with the fields of automobiles and home support robots that have similar issues. We added an explanation in the latter half of section 5.1.

5 Technology for accurately detecting risk

Question and Comment (Motoyuki Akamatsu)

You mention risk detection as an important technological issue, but risk decision that matches the user's risk recognition is extremely difficult as mentioned in section 5.1, and future R&D is essential. Therefore, the "risk detection" in section 3.5 is an explanation of the technology for obstacle detection (including dips). I think you need a description on the decision for the degree of risk and the decision to decelerate or stop.

Answer (Yutaka Satoh)

As you indicated, the technology for accurately detecting risks was insufficient, so we added the description in section 3.5 and newly added Fig. 8. For the analysis of level differences, we described that the detailed analysis of bumps and dips is difficult due to precision issues in the current system, and that we are separately working on a stereo image processing system using near-infrared pattern projection to solve this problem, and added a reference for this research.

6 Gesture detection

Question and Comment (Motoyuki Akamatsu)

The description about gesture detection is rather simple, and the range of application is not clear. Please provide a more detailed explanation for sensing and judgement on gestures and seating positions.

Answer (Yutaka Satoh)

As you mentioned, the description of gestures was insufficient so we added descriptions in section 3.6 and section 5.1. Specifically, we added the points: currently implemented are (1) the function to detect the abnormality of the seating position and (2) the function to detect the gestures of the arm; and we presupposed large motions since currently the gestures are determined by simple matching of quantized three-dimensional patterns. However, in actual fact, there is high expectation for gesture recognition from people who can move only parts of the body due to handicaps. Specifically, there is a request for recognizing the gesture of slightly moving the shoulder and we are working on it, but unlike the assumption of relatively large motion in the current technology, normal action and gesture motion cannot be separated and recognized at this point. Simple matching is insufficient, and we are considering a learning pattern-matching method. We added a description on this.