The Development of an Autonomous Field Transport Vehicle With an Active Vision System

Takashi Gomi and Koichi Ide
Applied AI Systems, Inc.,
340 March Road, Suite 600, Kanata, Ontario, Canada, K2K 2E4
{gomi, ide}@AAI.ca

Abstract: The development of prototype autonomous transport vehicles for use in construction sites is described. Such vehicles are expected to be in high demand in the near future in countries where aging of construction workers is becoming a serious problem. Two prototypes have been built so far and have undergone extensive testing. Vision systems used in the very rough operational environment of construction sites is discussed, as well as other aspects of the vehicle. A unique vision-based navigation operated well under extremely dynamic conditions. When completed in the near future, the project is expected to create a vehicle which would lessen the burden of workers in construction sites.

Keywords: autonomous robots, navigation, active vision

1. Introduction

There is a rapid and serious aging of civil and construction workers in industrialized countries. The phenomenon is already acute in Japan, where it is common to find workers in the field ranging in age from 50's to 70's. With this comes a lessening of the workers' physical and willingness to carry out tasks that demand extraneous physical efforts.

A great percentage of such physically demanding tasks involves hauling materials and tools for construction. The intention of the development project is to provide autonomous and semi-autonomous vehicles for transporting such items in actual construction sites. Since 1995, we have been developing a series of autonomous wheelchairs for eventual use by the handicapped and the aged[1,2] as well as autonomous office[3] and factory[4] vehicles. Through that development, we acquired knowledge and experience in building vehicles that can travel through poorly defined environments using various types of cues. One such type of cue is color markers of varied size and shape, either stationary or mobile. This technique is used in this development along with other navigation and manoeuver techniques.

2. The First Prototype

The first prototype vehicle named M-1 was built throughout 1997[5]. It was built on the framework of an electric scooter for senior citizens. It was fitted with a replaceable wooden framework for transporting rugged materials. The necessary electronics to support on-board intelligent navigation were installed. A set of active infrared sensors and a ccd camera were strategically installed around the wooden structure. The manually driven steering mechanism was motorized. The navigation unit responded to sensor signals and generated signals to the motor controllers for each of the two rear drive wheels and the steering motor. The on-board processing system consisted of a vision processor based on a digital signal processing (DSP) board that processed signals from a ccd color camera, and a 32-bit processor that handled other sensor inputs and the issuing of control signals to the motors. The building of the transport vehicle took some 5 months. M-1 was then tested extensively throughout the summer in and around our laboratory in Ottawa, Canada. In October 1997, M-1 was brought to an actual construction site on Awaji Island, Japan. Here it underwent a series of extensive testing in varied operational conditions.

The vehicle's navigation system was designed to follow a construction worker wearing a red safety vest commonly worn by workers at road and other construction sites. The red color area of the jacket was followed by the vehicle's single on-board camera fixed in the center of the 70 cm high and 65 cm wide front panel of the vehicle. The vision system was capable of processing about 12 frames per second using a TI C31 DSP chip. The small ccd camera was subjected to constant and often violent vibrations while the vehicle was in motion. The frequency and amplitudes of the vibration the camera was exposed to were up to ten or more cycles per second and maximum of 15 centimeters, respectively. Thus, it was not a stationary vision system as found in most vision-based mobile robot navigation, but an unstable image processing system with little or no continuity between frames of the incoming image. Since successive frames of the input image could be shifted as much as 15 cm at the camera, the latency of the motion generating complex of M-1 must play the role of integrator to smooth out the resulting over-all motions of the vehicle.

In the "Follow a Person" mode, the camera is constantly adjusted to place the center of the color mass in the center of the field. Whenever a shift which is greater than 6 pixels between these two centers is detected, the vision system issues a command to orient the body of the robot to align itself to what is found in the visual field. This, of course, cannot be achieved instantly because of the huge mass that lies between the visual frame and the body of the robot itself, plus the load it is carrying. Nevertheless, the command is issued as often as several times a second without any attempt to calculate appropriate turn angle for the alignment, as might be done in more conventional control systems. Despite this naive method for managing the robot's body, the emergent trajectory of M-1was very satisfactory.

Through the testing, the vehicle worked as envisioned for the most part, but was judged lacking in power. The less than sufficient power often resulted in uneven speed of the vehicle due to varying surface conditions commonly found at construction sites. The speed of M-1 was easily affected by unexpected encounters with obstructions such as rocks and holes resulting in varying distance between it and the worker leading it. The distance often grew too large for the vision system to keep the target marked, and this in turn resulted in the inability of M-1 to find the worker once it overcame the hindrance. Several attempts were made to adjust the software so that the problem could be resolved. In one such attempt, a non-linear power control scheme was introduced when M-1 encounters an obstruction and the on-board sensors detect slowing down of its speed. This strategy often allowed M-1 to successfully clear the obstruction without losing too much speed, but also resulted in unruly acceleration after overcoming a blockade, sometimes catching up and colliding with the worker leading the vehicle. Even when there was not a clear obstruction, the vehicle would sometimes slow down when faced with an inclination commonly found in many construction or civil engineering sites. This was a clear sign of a vehicle underpowered for a given application.

On-board active infrared (IR) sensors created another problem. Sunlight's wide spectrum easily covered the receptive frequency range of the IR sensors used onboard for detecting obstacles frontally facing the vehicle. This resulted in the saturation of the infrared receptors, causing them to generate maximum output regardless of the amplitude of the signal received. This created a situation in which M-1 was faced with a permanent phantom obstacle whenever its sensors are exposed to the sun. Sun visors of various shapes and sizes were tried on the sensors to prevent the sun's beam from directly shining into the receptor. This worked in some instances but failed in others, particularly when the sun's angle is low as in early morning or near sunset. Parameters of filter circuits within the IR sensors were also adjusted to increase the signal-to-noise ratio of the receptor. The effort resulted in some improvements, but again not enough to completely alleviate the problem. In the end, it was concluded that the use of IR sensors in this mode for application in outdoor vehicles is not feasible.

3. The Second Prototype

3.1 Hardware structure of the prototype

It was decided at the end of 1997 that a new prototype vehicle with larger motors and power system called M-2 was to be constructed. The development begun early in 1998 and lasted to the end of September 1998. M-2, shown in Figure 3.1, was ready for first testing by June 1998. A small-size electrical golf cart chassis was used

Figure 3.1 The second prototype, M-2

as a base. Again, the superstructure of the original cart was replaced with a wooden carrying box. Sturdy hard rubber bumpers were added to the front and the back of the vehicle. Four clearly visible emergency stop switches were added for safety. Larger motors, 400 watts each, that came with the cart were judged sufficient. As in M-1, the steering mechanism was motorized.

The hardware structure of M-2, shown in Figure 3.2, is basically the same as that of M-1, but with some changes and additions. Sonars were chosen to replace active IR sensors for avoiding collisions. The MC68332 board deals with all the non-vision processing of the on-board system. This includes processing of signals from sensors such as bumpers and sonars, and remote command inputs such as the newly added radio and voice commands. It also generates drive signals to controllers for the two wheel drive motors, the steering motor, and the camera's pan motor, which was added during testing. Most of the signals in and out of this main processing board are handled through a specially designed interface board. This board also manages the low-level protocols for the IR communication system which was newly added to M-2. Both voice input/output and radio command subsystems are only for preliminary testing of the concept on M-2, and not extensively tested during this phase of the development.

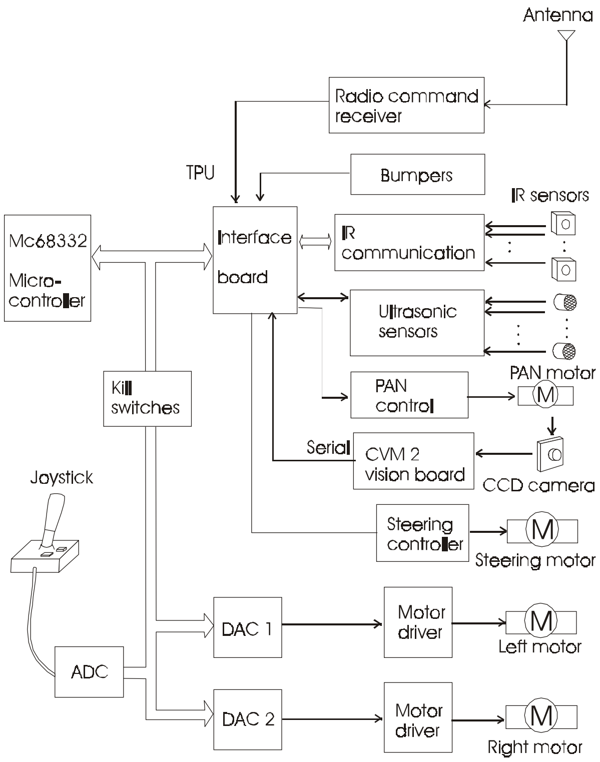

Figure 3.2 Hardware structure of M-2

The IR communication subsystem is introduced to expand the operational versatility and autonomy of the vehicle. It consists of one or more IR command modules that are placed on a safety pylon (or other appropriate structure in the field), and an on-board IR receiver. The IR command module, which is battery driven, can be set to transmit one of six predefined commands to its immediate vicinity. The commands are: "Turn left", "Turn right", "Circle left", "Circle right", "Stop", and "Pause until". The default command is "Continue". Using this communication, the flexibility of M-2' s navigational capability is greatly increased. In addition to the "Follow a person" mode of M-1, M-2 can now operate fully autonomously. This was achieved by adding software for visually tracking a series of objects such as above mentioned safety pylons. By marking a course M-2 should follow by a series of red plastic safety pylons placed every 10 meters or so, the vehicle would follow the trail as long as the pylons were found, or until instructed otherwise by an IR commander. Since landscapes at construction or civil engineering sites tend to change almost daily, this simple, flexible, and easy to do manner of setting up and adjusting the course should help the users.

3.2 Software structure of the prototype

The software structure of M-2 is shown in Figure 3.3. It consists of a number of *competence* (or "component behavior") modules and a simple priority arbitration mechanism. Not all *competence* modules are shown in the diagram due to space limitations. With software improvements, the vision system now processes about 20 frames per second compared to the 14 of M-1' s vision system. The signal from the vision system is fed to a set of *competencies* that handle navigation and manoeuvring of M-2. ***Follow*** generates signals to motors that track the worker leading the vehicle. Both ***Right Pass*** and ***Left Pass*** generate drive signal to motors so that the vehicle follows either the right or the left of the pylons, respectively. Each of the instructions issued through the IR communication channel is supported by an on-board *competence*. Thus, a "Stop" signal detected by the IR receiver is implemented by ***Stop*** *competence* to immediately halt the vehicle. Similarly, turns and "circle" signals are executed by corresponding *competencies.* Note that the outputs from IR-instructed *competencies* are given a higher priority than those from vision-driven *competencies*. This assures that vision-based navigation can be interrupted at any time by instructions sent through IR commanders. Figure 3.3 also implies that the detection by one of the ultrasonic sensors of an approaching obstruction are treated at an

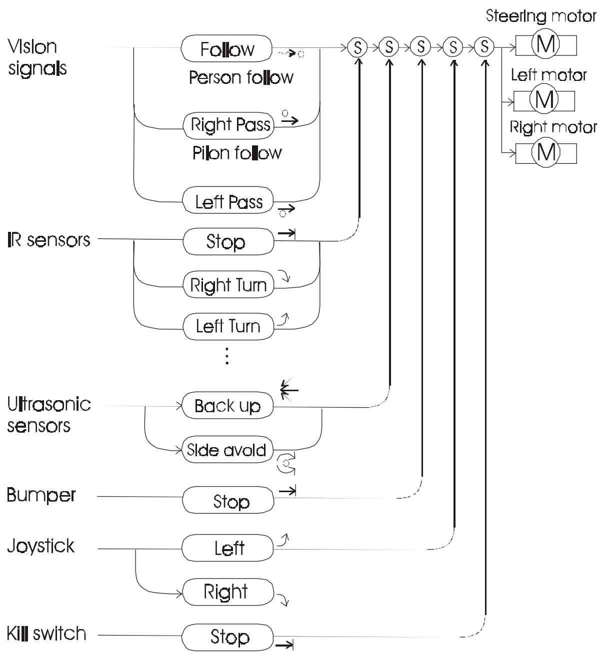

Figure 3.3 Software structure of M-2

even higher priority. This in turn is superceded by ***Stop*** *competence* invoked by a signal from a bumper. Joystick-driven *competencies* have yet higher priority, allowing M-2 to be controlled by hand under any situation. Not all *competencies* are implemented as a software-driven process. They can be implemented by a multitudes of methods either in hardware or software [6]. Joystick *competencies* are implemented as electronic circuits that convert voltage changes at the joystick into driving signals sent to the motors. This allows the manoeuver of M-2 even when the processor is down. Pressing one of the kill switches immediately stops all operation of the vehicle by blocking inputs to the motor drivers.

Notice that the execution of *competences* does not imply an execution of a prearranged plan of any sort. All *competencies* are invoked solely by external stimuli. There is no centralized control in the system. Governing the output of *competencies* by a strict fixed priority scheme does not result in inanimate regimented motions as often seen in mobile robots driven by a centralized and/or hierarchical control structure. The second by second trajectory of M-2 emerges as a result of self organization arising from asynchronous, parallel, and involuntary execution of multiple *competencies*. Despite the lack of plan and centralized control, the vehicle manages to follow a desired path while avoiding obstacles and dealing with other interruptions as necessary. In fact, M-2 exhibits very fast responses to changing situations in its environment, required in vehicles operating in real world applications.

3.3 The vision system

The strategy with which M-2' s vision system operates has evolved through a large number of field tests. The motor control mechanism for steering looks at the current frame at 64 Hz and sends out a set of new motor signals. The mass of the vehicle and the mechanical impedance of its steering mechanism jointly work as an integrator and alter the course of actions of M-2 at a much slower pace. When in the "Follow a person" mode, the vision system always tries to maintain the worker's red safety jacket in the horizontal center of the view regardless of the heading of the vehicle's body. This is achieved by sending correction signals to the pan motor at 64Hz. The body then tries to align itself with the orientation of the camera's optical axis.

Although a clear color is chosen for markers, their actual appearance in the camera's field is not always

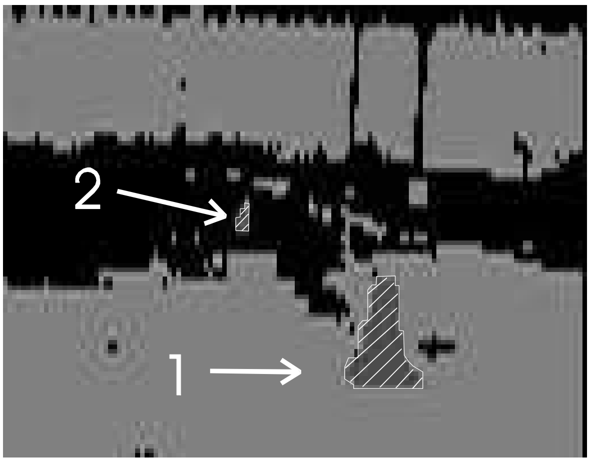

Figure 3.4 Two pylons in frontal lighting condition

clear. Figure 3.4 is a more fortunate case where the sun is behind the camera and two pylons are clearly located after a color segmentation process, during which areas with similar colors are gathered together. Once the target color markers are identified, other color areas are turned into levels of grey. Although simply shown as shaded areas in this black and white reproduction, pylons actually appear as distinct orange patches in the middle of grey patterns. Figure 3.5 is a case where the camera is facing the sun. Two consecutive pylons (1 and 2) appear in the path of the vehicle. Only the shaded patches around pylon 1 appear as orange, and the rest as grey levels. This forces M-2 to proceed cautiously according to the preset navigation mode (in this case, "Pass right of pylon"), until the vision system finds pylon 2, or a certain distance is covered unsuccessfully. Note that all other sensors are active at higher priorities through this time, meaning any impromptu encounters with obstacles would be dealt with quickly. Following early testing in which this problem of limited vision was

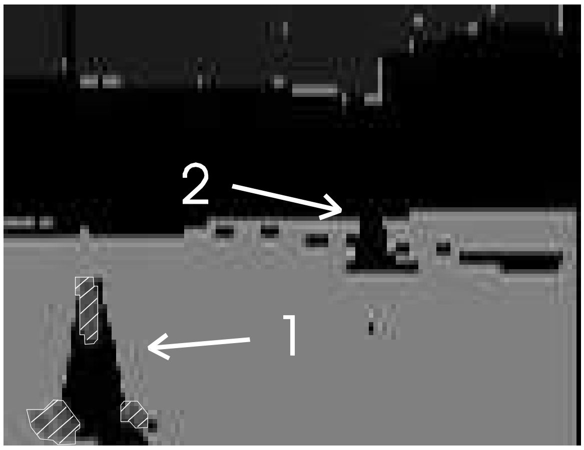

Figure 3.5 Two pylons in reverse light

found, we added a simple panning scheme to the camera. This allowed the camera to spot color markers which the fixed camera could easily overlook. The importance of saccadic eye movements in active vision are emphasized elsewhere [7,8]. Our example demonstrates that it is also effective in vision-based navigation in which capturing of a frame is highly dynamic, unstable, and often sporadic.

The camera's field is 256 by 100 pixels, but is reduced to 128 by 50 pixels on the DSP. If this field does not contain a marker after color segmentation, the panning mechanism sweeps between +-60 degrees and +60 degrees continuously. If a marker is captured and it is less than 50 pixels wide (out of 128), the camera repeatedly zooms up 200% around the marker and centers it horizontally until the condition is satisfied. If the marker in the frame is more than 70 pixels wide, the camera zooms down. Once the marker is captured, in "Follow a person" mode, the camera constantly pans in small increments so that the marker stays in the horizontal center of the field regardless of the second-to-second orientation of M-2's body. In "Follow pylons" mode, upon capturing the marker, the vision system: (1) tries to keep it in the horizontal center of the field if the total area of the marker is less than 64 pixels; (2) pans the camera to keeps the center of the marker at 16 pixels from the horizontal edge of the frame, if the area of the marker is between 64 and 255 pixels; (3) ignores the marker's relative position in the field and proceeds straight forward, if its size exceeds 255. These parameters were chosen after a large number of tests in the field. They effectively allows M-2 to travel from one pylon to the next along a smooth trajectory under most lighting conditions.

More demanding cases arise when the sun's angle is low as in early morning or late afternoon. The shift of sunlight's color spectrum to red due to its longer path through the atmosphere results in a reddish scene. Figure 3.6 shows a view through M-2' s vision system at a bright sunset. M-2 is placed on a brick covered sidewalk which is bordered on its left by a line of shrubs and pavement on the right. Some trees appear ahead in the view. Regardless, the entire scene appears as a yellowish red pattern, completely mixing everything including two red pylons and trees in the background. The color segmentation algorithm barely restores some patches of red around pylon 1, but nothing on or around the second pylon. This forces M-2 to repeat the cautious inter-pylon moves mentioned above. It is necessary to have a better color compensation scheme to match the wider-than-expected range of changes in conditions created by various states of the sunlight.

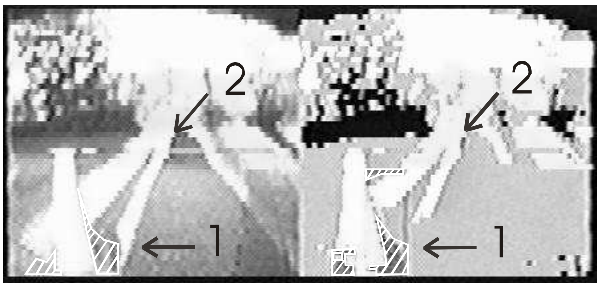

Figure 3.6 Two pylons at sunset

4. Tests Using the Second Prototype

M-2' s testing begun in June 1998, first indoor at our laboratory, and then outside. It was driven extensively over a grassed landscape, asphalt surfaces, and finally over an actual construction site made up of small bushes, soil, gravel, and rocks. The site included very rough terrain. The vehicle was first tested extensively in its "Follow a person" mode, and then in "Follow pylons" mode. In October 1998, M-2 was sent to Japan for field tests. A series of tests was carried out in the remote town of Kosaka, Akita Prefecture. This session included testing at an actual construction site, at a road construction site, and a large cabbage patch.

Towards the end of the tests the vehicle performed well in the field and actual construction sites. But through this test period several issues were brought up, most of which we did not anticipate. A few of the biggest problems were not even related to the autonomous running of the vehicle. They were burnt motor controllers, worn gears, and short-lived batteries. For various reasons, altogether 4 sets of different controllers were destroyed and replaced during the development of M-2. Each time the project was significantly delayed. Batteries required replacing prematurely due to malfunctioning of the automated battery charger. All these indicate the importance of engineering issues when developing real world applications.

Beyond this, the revision of the vision system took most of the development time. Once good vision performance was obtained, both the "Follow a person" mode and the "Follow pylons" mode worked satisfactorily most of the time. Since the original golf cart was designed for relatively light use, when loaded in actual construction sites, it showed the sign of weakness. It was felt that more powerful motors were desirable to smooth out the runs. Batteries with higher capacity would also be desirable, as we saw occasions when batteries (55Ah x 2) were exhausted sooner than anticipated. We learned the hard way that the selection of the right base vehicle is crucial.

Additional suggested improvements include the following:

(1) ability to follow a person with more even distance,
(2) higher speed (of up to 2.5 meters per second) in open spaces,
(3) additional operational modes (governed through the IR communication channel),
(4) stable vision system particularly against changes in lighting conditions,
(5) improved cooling system for the motors,
(6) built-in charging system,
(7) better suspension,
(8) simpler user interface,
(9) higher bottom clearance,
(10) more sensitive bump sensors,
(11) improved box design, and
(12) better mounting of the steering motor.

Ultrasonic sensors (sonars) worked well at construction sites. They caused less problems outdoors from secondary and tertiary reflections as obstacles are far apart and the space generally not crowded. The IR communication channel initially had a problem of sun light dwarfing the received signal. This was overcome by increasing the number of LED's on a transmission tower from 6 to 16, receding the on-board IR receptors to avoid direct sun light, and by adding a lens to the receptors. After these modifications, the system worked well up to a distance of several meters from the initial 50 cm. In the end, stable outdoor communications made the fully autonomous operation possible.

5. Results of the Development

The IR communication subsystem used in M-2 proved very effective. The facility became a useful channel for issuing instructions to the vehicle. M-2' s active vision system makes the vehicle follow a series of markers, which are casually and imprecisely placed in rugged fields. The IR communication supplemented the operation by passing crucial bits of information whenever the vision guided vehicle needed it. The two worked complementary to each other. Additional commands beyond the current 6 could be easily added to make the repertoire of operational modes richer.

There are still problems to be solved for the vehicle to be used in daily operation. The dynamic range of the on-board camera needs to be widened so that it can function in more demanding lighting conditions. It is concluded after the evaluation of M-2's test results that a faster vision processor would have to be developed to push the vehicle's performance. The improved vision processor would be incorporated in a new prototype vehicle (M-3) to be constructed later.

M-3, when implemented, will not only have an improved vision system but also incorporate most of the "desirables" discussed above. Both the radio and the voice communication subsystems will be activated and tested extensively in the field. One idea we are contemplating is to combine the two so that the vehicle could be managed remotely using voice commands. For example, a worker could retrieve the vehicle from the other corner of the site by radioing it to "Come back to base", or guide the vehicle by talking to it via radio to "Go another 25 meters".

With or without such functional improvements, there is no doubt the vehicle must undergo many more tests. So far, the combined M-1 and M-2 passed the scrutiny of some 500 hours of indoor and field tests and most problems discovered during the tests were immediately fed back to the development team and corrected quickly. Then, a series of more formal performance and reliability tests would be mandatory before one could consider a public version of the vehicle with sufficient field-worthiness. For instance, the whole area of water resistance must be addressed as the current prototypes have only nominal protection against rain and other water invasions. Nevertheless, the prototypes more than amply demonstrated the utility of autonomous and semi-autonomous vehicles with features and capability of M-2 in average construction and civil engineering sites. While the prototypes are tested in location, they captured the eyes of workers, supervisors and operators of construction businesses alike as a field tool with a high potential for success.

Acknowledgements

Takashi Gomi created the functional and methodological designs of the two prototypes and supervised the execution of the project. Koichi Ide was responsible for the detailed design and implementation of the prototypes. Patrick Maheral assisted Ide in designing and implementing the vision systems used in both prototypes. The bulk of the implementation of the vehicles was carried out by a number of co-op students in engineering departments of University of Waterloo in Canada between January 1997 and October 1998. Of the co-op students, Euhan Chong worked extensively on M-1 in circuit implementation and testing, and also testing of that prototype in a construction site in Japan. He also worked on the construction of the body of M-2. Robert McConnell, Landy Toth, Alak Ghosh mostly worked on M-2 in its construction and testing. Brian Koh of the National University of Singapore assisted the building and testing of M-2 in Canada. Akira Matsumoto of Fukui University in Japan and Fumio Tsukada of AAI Japan, Inc. assisted in the testing of M-2 in Japan. The development team is grateful to Matozaki Construction Co. Ltd. of Awaji, Japan for their generous support.

References

1) T. Gomi and K. Ide, The Development of an Intelligent Wheelchair, Intelligent Vehicles'96 (IV'96), 1996.

2) T. Gomi and A. Griffith, Developing Intelligent Wheelchairs for the Handicapped, Lecture Notes in AI: Assistive Technology and Artificial Intelligence (Springer Verlag), 1998.

3) T. Gomi *et al.*, *Vision-based Navigation for an Office Messenger Robot*, Chapter in Intelligent Robots and Systems, ed. V Graefe, (Elsevier Science, 1995).

4) T. Gomi, *Fully Autonomous Ground Vehicle (FAGV) for Industrial Applications*, Proceedings of International Conference on Robotics, Vision, and Parallel Processing For Industrial Automation (ROVPIA'96) (Malaysia, November 1996).

5) E. Chong, "Performance and Analysis of M-1 Flexible Autonomous Ground Vehicle". University of Waterloo work term report, January, 1998.

6) T. Gomi, Impact of Non-Cartesianism on Software Engineering, Chapter in Computational Intelligence in Software Engineering, edited by W. Pedrycz and J.F. Peters, World Scientific, 1998.

7) D.H. Ballard, Reference Frames for Animate Vision, Proceedings of the International Joint Conference on Artificial Intelligence, Minneapolis, MN, 1989.

8) R.A. Brooks et al, Alternative Essences of Intelligence, Proceedings of the Fifteenth National Conference on Artificial Intelligence, Madison, Wisconsin, 1998.