Real-Time Visual Behaviours for Navigating a Mobile Robot

Gordon Cheng and Alexander Zelinsky
Department of Systems Engineering
Research School of Information Sciences and Engineering
Australian National University
Canberra, ACT 0200.
Australia
email: Alex.Zelinsky@anu.edu.au

Abstract

In this paper we present an approach for using vision as the primary source of sensing to guide a mobile robot in an unknown environment. We define a set of primitive visual behaviours for navigating a mobile robot in real-time. By combining such behaviours with a purposive map, our mobile robot exhibits a goal seeking behaviour. We present a fast segmentation technique for vision processing. This processing technique is used by different behaviours to produce an overall competent behaviour in our Yamabico robot. Experimental results show that our robot can navigate competently in dynamic indoor environments.

1. Introduction

For robots to work harmoniously within an integrated environment with humans the robots must have navigational skills. In a human only environment, human may tolerate collisions with one another if they did not cause much pain. This level of tolerance may or may not apply to robot-human environments. Humans expect that robots will be able to operate and navigate in their environments without collisions or interference.

The use of vision in mobile robotics has become wide spread. Many researchers have applied vision to collision avoidance such as Horswill [1] with the “Polly” robot and Gomi *et al.* [2] with their “Office Messenger Robot”. Vision has been used in different ways for example door opening was reported by Nagatani *et al.* [3] and car assembly by Kimura *et al.* [4]. Traditionally most techniques follow the conventional AI ideas of building a model of the environment together with having a thorough understanding of such an environment in order to perform navigational planning [7]. These techniques become constrained by the computational power of the machine that is being used. For obstacle avoidance most methods try to determine the position of each obstacles in order to move around them. We agree with Horswill’s [1] approach in which obstacle avoidance is based on searching for features that are physically grounded e.g. carpet. We refer to this as free space robot motion.

Humans have a good sense of direction, and can navigate without using precise coordinates for localisation [16]. Navigation is done relative to landmarks [17]. Our research goal is to develop a robot navigation system that is similar to the human system.

Humans navigate using vision very successfully. This is something we take for granted. Consider a young child that has acquired the ability to crawl. The child wanders around the floor investigating its new environment. The child’s motor skills are well developed for manoeuvring on the floor. The child learns to avoid furniture and moving objects such as people. The young baby may wander around the floor until it finds a toy it likes. To give a child a purposeful desire to move in a particular direction a parent could show the child their favourite toy. The child responds to this stimulus by moving toward the toy. At the same time the child will avoid any obstacles in its way. From this example we see that the child has leant basic navigation skills. Through this development process we can established that all the essential behaviours were induced [9]. First the child gained motor skills, then with the assistance of vision is able to navigate. From our point of view three primary visual-based behaviours are needed; goal seeking, obstacle avoidance and collision avoidance. These we believe are the fundamental behaviours for a mobile robot to achieve a much more natural acceptance in the human world.

The aim of our research is to provide a mobile robot with sufficient visual behaviours to navigated freely in an unknown indoor environment. At the same time a purposeful desire to pursue a goal if it is provided to the mobile robot. Such a robot may be used as a scout for a larger group of less efficient robots. It could be sent ahead to investigate the environment and return with the shortest/safest path for their mission.

Initially we have developed this system using an off-board vision system. However, after development of the visual behaviours we plan to design a detailed specification for an onboard vision processing system for our mobile robot.

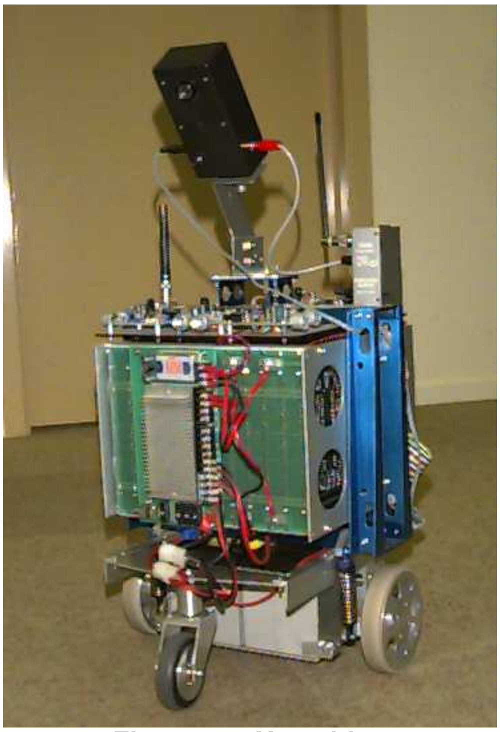

Figure 1. Yamabico

2. Our System

The robot we are using is a Yamabico mobile robot that was developed at the University of Tsukuba by Yuta *et al.* [11]. It has a distributed system with separate modules for locomotion, ultrasonic sensing plus a master module for coordinating the processors. To our Yamabico mobile robot we have added a small size video camera with a video transmitter and radio modem, as shown in Figure 1. The video transmitter is used to transfer the video signal from the robot's video camera into the vision system for processing. The radio modem on the robot is used to receive for information that was processed by the vision system. An overview of our system is shown in Figure 2. The robot transmits the video frames to the vision system, which processes the video signal. The vision system then sends the sensory data through to the communication server. The communication server passes information back to the robot. This forms a complete loop. Further details of each of these components are discussed in the following sections.

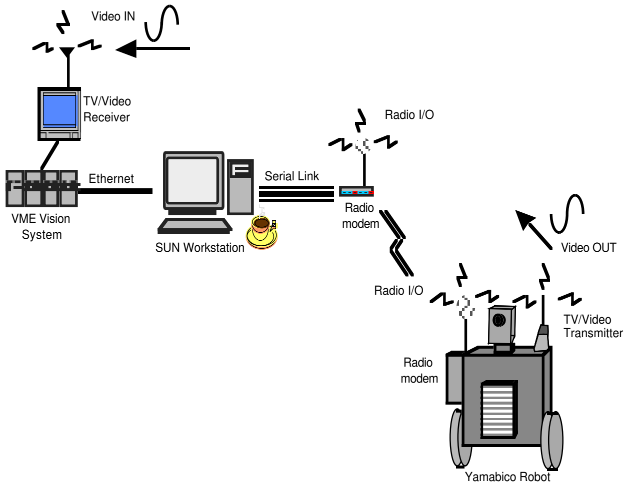

Figure 2. Configuration of our system

2.1 Real-Time Vision System

The vision system we are using is a MEP Tracking Vision system that is produced by Fujitsu Japan. It is designed for real-time tracking of templates. It can track up to 72 gray-scale templates at frame rate. The system is powered by a VME-bus. It is equipped with one video module, one tracking module and a MC68040 processor module running the VxWorks® operating system.

The purpose of the video module is to take video signals, which is then processed by the tracking module. This processing is done through the use of three video memories(VRAM) in the system, which are shared by both modules. Tracking is done by using one VRAM as an buffer for live video, and another VRAM for the storage of templates. Each template can be of a size of either 8x8 or 16x16, with a magnification of up to 4 times a template. A distortion between the two VRAM memories is calculated by the tracking module at real-time at a given location of the live video image. This will perform a cross-correlation match between the live image to a given template, for determining the confident of the match. Returning a low distortion from the tracking module means that the template is being closely matched and tracked.

The role of the vision system is to process all the video signals from the robot's camera. There are other projects in our lab that also utilise this vision system. Refer to Jung[5] and Heinzmann[6]. As a prototyping system for our visual behaviours this off-board vision system is sufficient for our use. Once we have determined the exact specifications for our visual behaviours, we plans to develop an onboard vision system.

2.2 The Communication Server

The communication server we are using is a SUN-Sparc workstation running the Solaris operating system on an ethernet network, equipped with a radio modem. The primary usage of this server is to pass information from the vision system to the robot. This is done via the ethernet network from the vision system. The server passes on sensory data from the vision system over the radio modem to the robot. The communication server collects then processes the debugging information that has been sent between the vision system and the robot. The server can also perform functions such as drawing the robot's execution path. Any additional processing does not increase the computational load on the robot and on the vision system.

2.3 The camera

We are using a light 38 grams CCD camera mounted on top of our mobile robot at an angle facing the floor The camera we are using has a viewing angle of 85 degrees, which provides a minimum amount of distortion. Therefore a simple projection model can be used to estimate the floor space. This estimation model is used for transforming the screen pixel of the image to the floor coordinates. This estimate can be used by each of the visual behaviours; goal seeking, obstacle avoidance and the collision avoidance behaviour.

The coordinate transformation is calculated by measuring four distances of the floor, as illustrated in Figure 3 and Figure 4. The first distance is the height of the camera to the floor, $height_y$. The second distance is the nearest view point of the camera, $blind_y$. The third distance is the furthest visible point to the top view of the camera, $blind_y$ plus $length_y$. The third distance is the furthest point to the left/right of the camera view, $length_x$. With these measurements we can calculate the three angles required, using Equations (1), (2) and (3). From the angles we can determine the angular ratio in both the x and the y directions. Figure 3 shows the side view of the robot. Figure 4 shows the top view of the robot.

$$\alpha = \tan^{-1}\left(\frac{height_y}{blind_y}\right) \qquad (1)$$

$$\theta = \tan^{-1}\left(\frac{height_y}{blind_y + length_y}\right) \qquad (2)$$

$$\beta = \tan^{-1}\left(\frac{blind_y + length_y}{length_x}\right) \qquad (3)$$

The following equation calculates the transformation of the y^{th} pixel to the vertical distance on the floor.

$$y = \frac{height_y}{\tan\left(\theta + \frac{pixel_y(\alpha - \theta)}{screen_y}\right)} + \frac{robot_y}{2} + blind_y \qquad (4)$$

where $pixel_y$ is the y^{th} pixel of the screen. $robot_y$ the length of the robot in the y direction.

The following equation calculates the transformation of the x^{th} pixel to the horizontal distance on the floor.

$$x = \tan\left(\frac{\beta(1 - 2pixel_x)}{screen_x}\right) \times y \qquad (5)$$

where $pixel_x$ the x^{th} pixel of the screen.

From the Equations (4) and (5). we can calculate the floor space (x, y, θ) for the robot to travel towards from the corresponding point in the image space.

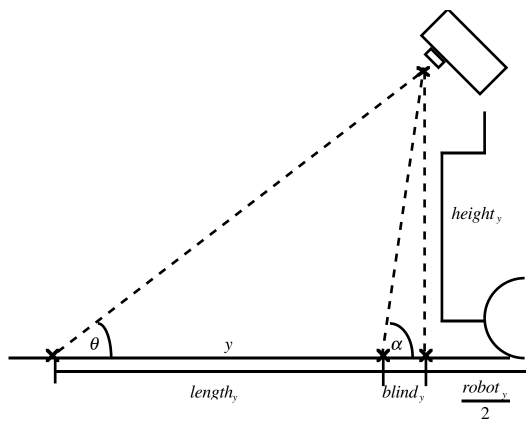

Figure 3. Side view

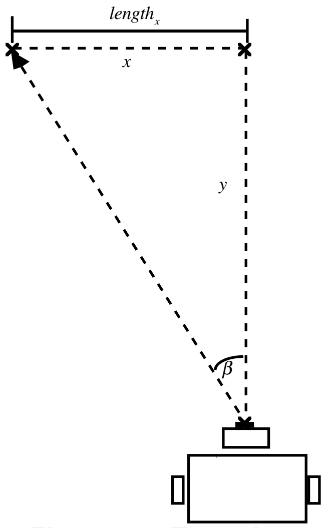

Figure 4. Top view

2.4 Visual Sensory processing

The data processing on our vision system is broken up into two tasks shown in Figure 5. One task is dedicated to processing the video signal while the other process calculates the actuator commands to be passed to the robot. The video signal processing tasks categorises templates. For example in the Free Space Motion behaviour this task determines if a block of pixel (template) is either carpet or an obstacle. Further details of the segmentation technique is discussed later. All video signal processing is done at frame rate, which in our system is 30 Hz.

The filter task performs filtration of the actuator commands it has collected from the output of the vision process at a set time interval. This filter uses an histogram of all the robot actuator commands that was collected in an given period of time. It determines which is the most commands which needs to be sent back to the robot's actuators. This process can run up to 30 Hz.

A common data queue is used to share the processed data between the vision process and the filtering process. Once an actuator command reaches the robot it is sent to the Spur [15] locomotion system of the Yamabico robot. The Spur systems accepts a (x, y, θ) trajectory along which to head. The locomotion sub-system insures the robot tracks along this desired trajectory. In the case of the Free Space Motion behaviour the system commands the robot to move along a trajectory into the largest free space floor area.

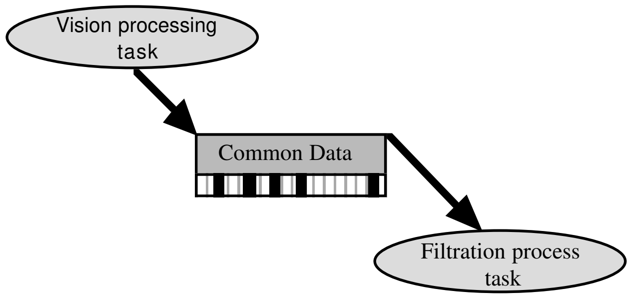

Figure 5. Visual Sensory processing

2.5 Blind spot

With the current system configuration a 0.4m dead zone exists below the camera which could not be observed by the robot. We believed that a secondary sensor can compensate for this dead-zone. Such as a ultrasonic or a bump sensor. However, our robot navigation system is quite robust, and the need for these secondary sensor is minimal. At this stage we plan to reply solely on visual sensor.

2.6 Safety and Speed

In our laboratory with many researchers it would be frightening to have a moving object running at up to 1 m/sec colliding with anybody. Therefore a safety constrain was put into the design. The robot keeps a clearance of 0.5 metres away from any obstacle. We also control the velocity of the robot while it is turning. The top speed of the robot is determined by the visual depth of the image, using Equation (6). The velocity of the robot while turning is determined by Equation (7). From this equation we can see that the robot will travel at its maximum speed only if it is moving straight, that is when $\beta = 0$ degrees.

$$top_{speed} = \max_{speed} * \frac{y}{length_y + blind_y} \qquad (6)$$

where $\max_{speed}$ is the maximum physically possible speed of the robot.

$$v = top_{speed} - \left| \frac{top_{speed} * \beta}{\phi} \right| \qquad (7)$$

3. Visual Behaviours

Our robot possesses a set of primitive behaviours, which are goal seeking, obstacle avoidance and collision avoidance. A robot agent must also possess a motivation

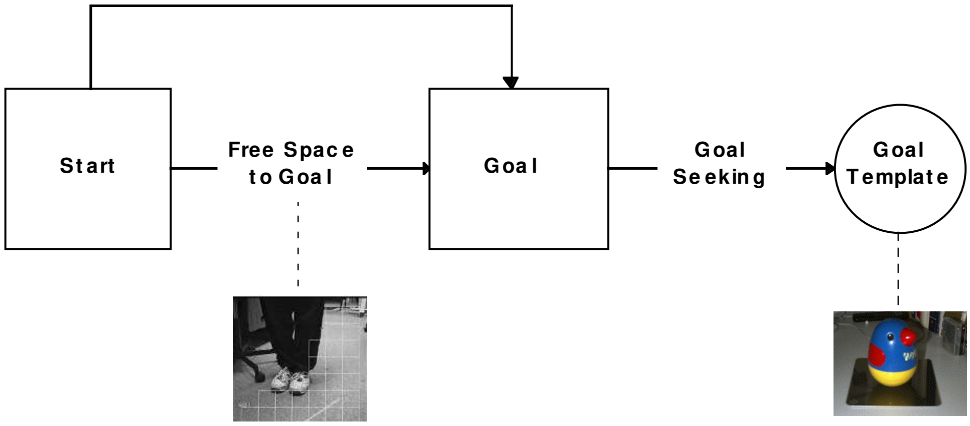

Figure 6. Purposive Map

and desire in order to accomplish purposeful tasks. In Zelinsky *et al.* [14] the concept of a Purposive Map (PM) was introduced. The PM is used to coordinate a robot's behaviours, and also provide a mechanism for resolving conflicts between behaviours. In this work Zelinsky *et. al.* [14] using sonar sensing for navigation. From multiple behaviours running concurrently a potential-like behaviour emerged. In our system, the PM is also used as a resource to coordinate the arbitration of our behaviours. New higher level navigation behaviours emerge from the concurrent running of primitive visual behaviours.

3.1 Free Space Motion

This behaviour produces obstacle avoidance by searching at frame rate for free space in each frame of the video signal. This is done by segmenting each frame into a grid of cells in which each cell is compared with a stored template of carpet. To reduce the problem of spatial depth information from the video imagery, a different template is used for each row of the grid. A grid of 8x7 cells was used. In this behaviour only the bottom 5 rows of the image are used, the remaining 2 rows are reserved for further study in landmarks acquisition. For each of these 5 rows we determine where free space can be found, from this we determine a avoidance vector that will take the robot into free space. The robot always heads for the top of the video image. Hence we determine a vector that will guide the robot to the highest free space in the image. This causes the robot to move away from obstacles. As a consequence of the free space move behaviour obstacle avoidance emerges. In the case of empty carpet the robot will keep moving straight. If an obstacle appears the robot will veer away from it.

Figure 7 shows a frame of the vision system performing a search for free space. This is shown in the figure by a white squares for each template that matches a carpet. The white line show the avoidance vector to which the robot should turn. The computation of this behaviour at frame rate results in a robust behaviour, even with noisy sensor data. Our robot can roam around our laboratory avoiding both stationary and moving obstacles.

Figure 7. Free Space Motion

3.2 Goal Seeking

While the robot is travelling through the environment it could be given the purpose of finding a goal, and then using this for navigational purposes. This is done by providing the robot with goal templates (or landmarks). In our experiments we sometime use a "Toy chicken" and a "Tennis ball" as our landmarks. These are shown in Figure 8. The vision system continuously searches its input for the landmark template. If a landmark template is detected, the robot can execute several actions. The robot can exhibit visual servoing, the location of the landmark template in the image is used as a goal for the robot locomotion system. This behaviour is useful for tasks such as pushing and picking up objects. If at any time a match is found a motion vector toward the "Toy chicken" is sent to the robot. This will guide the robot towards the goal. This is a visual servoing behaviour. Our robot can execute visual servoing behaviour to an accuracy of several centimetres. This goal seeking behaviour also serves as a cue for other robot behaviours. A tennis ball could mean turn left and go straight, while two tennis balls could mean turn right etc.

In the current system, templates are stored and pre-loaded before each mission. For further investigation, we are undertaking dynamic acquisition of nature landmarks as templates. Using environmental features as visual cue for navigating, techniques such as combining edge-detection and template matching is currently under review. Edge-detection has been used successfully for robot navigation dues by Ishii et al.[20] and template matching is a simple and effective approach to real-time visual processing.

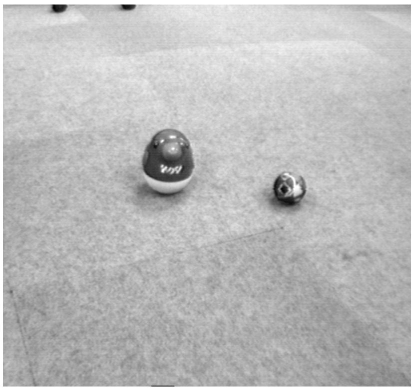

Figure 8. Goal Templates

3.3 Collision Avoidance

The Collision avoidance behaviour alone is very important, for a quick moving robot in unknown environments. As a matter of safety we have built a behaviour based on unconditional reflexes[5]. This behaviour overrides other behaviours. From the grid segmentation we can determine if an obstacle is too close, hence no free space can be found. The robot stops and then starts to spin on the spot. The robot spins until free space is sighted by the vision system. This causes the robot to stop spinning and then move forward.

As a further study we plan to experiment with the collision avoidance behaviour in the manner of conditional reflexes[5], that is acquired from past history of the robot. The robot leans to associate vision inputs with bump sensor data. Similar work has been done by Pfeifer [19] using other sensors.

3.4 Move to goal

This behaviour guides the robot to the direction of the goal. In the PM a landmark can be included together with information about its location of a goal. For example, that a goal is about 6 meters away in the N-E direction. This provides the robot with an internal representation of the goal. This is similar to the idea of McFarland *et al.* [18] of goal-directed behaviour. This provides the robot with a desire to move towards a goal. The behaviour produces a motion vector that is fed into the robot's locomotion system resulting in motion towards the goal. Combining this behaviour with the goal seeking behaviour results in a behaviour which moves the robot between landmarks. Our robot can navigate by using visual landmarks as cues to robot actions. Our robot can execute the type of mission shown in Figure 9. The navigation between the landmarks are physically grounded to the visual behaviours. The position of the landmarks can be changed without affecting the system. Thus yielding a robust system.

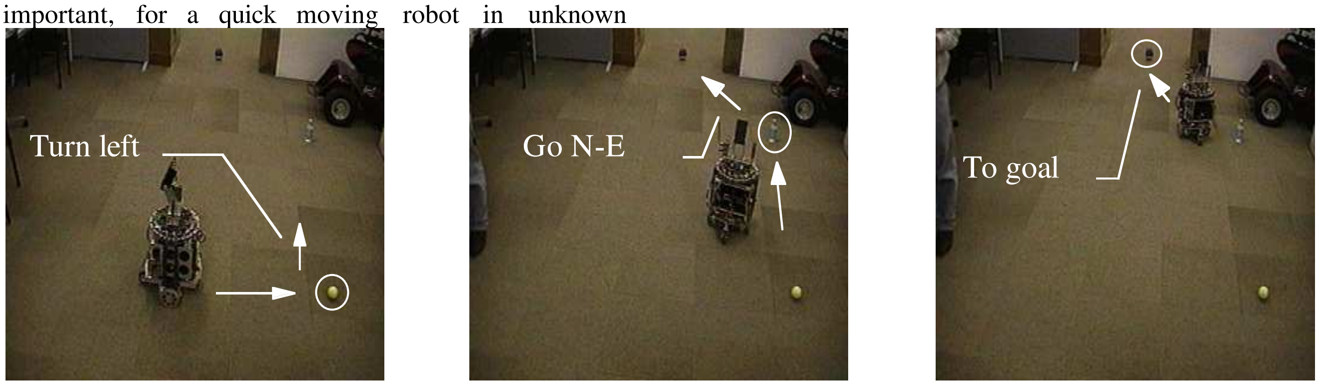

Figure 9. Mission

Figure 10. Local minimas

3.5 Emergent behaviour

A new visual behaviour emerges from the combination of the Move to Goal behaviour, with the three primitive visual behaviours; Free Space Motion, Goal Seeking and Collision Avoidance. This results in a behaviour which moves towards and seeks a goal while avoiding collisions with obstacles. This is similar to our earlier example of a child playing with a toy.

In the earlier work reported by Zelinsky et al. [14]. Competition between behaviours can arise which results in stagnation or oscillation between behaviours. The urge to move towards the goal is matched by desire to avoid obstacles. To solve this problem the PM is modified to include state information about behaviours. The earlier work concerning the PM used sonar sensing with potential fields navigation. This introduced two types of stagnation points(or local minima), free space minima and goal is blocked minima. Refer to Figure 10

Using our visual behaviours with free space motion and move to goal the free space local minima problem doesn't occur. However, we must deal with the goal is blocked minima.

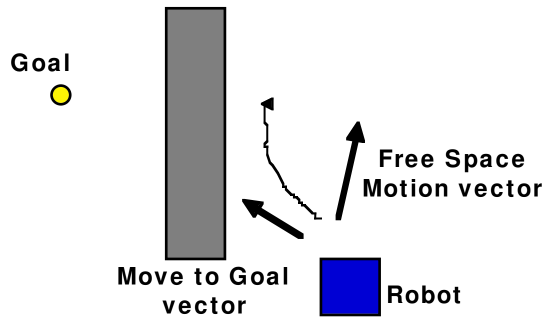

Figure 11. Emergence of contour following

By combining the free space motion and move to goal behaviours contour following of obstacles emerges. Refer to Figure 11. If the obstacle blocking the goal is not significant in size the robot will be easily able to avoid it and reach the goal. However, if the obstacle is larger a oscillation like behaviour will occur, due to the competition between the free space motion and the move to goal behaviours. We solve this problem by using the PM to store state information, in which the robot by trial and error discovers the correct persistence to give the free space motion.

4. Conclusion

We have implemented the visual behaviours that were described here on a Yamabico mobile robot with an off-board tracking vision system. We have implemented primitive visual behaviours; goal seeking, obstacle avoidance and collision avoidance. From these primitive behaviours emerge purposeful behaviours.

The current configuration, the robot can travel up to a speed of 300 mm/s. Figure 13 illustrates our robot navigating in our lab avoiding obstacles.

Acknowledgments

The research presented here is funded by a grant from the Australian Research Council and supported by Wind River Systems supplier of the VxWorks® operating system.

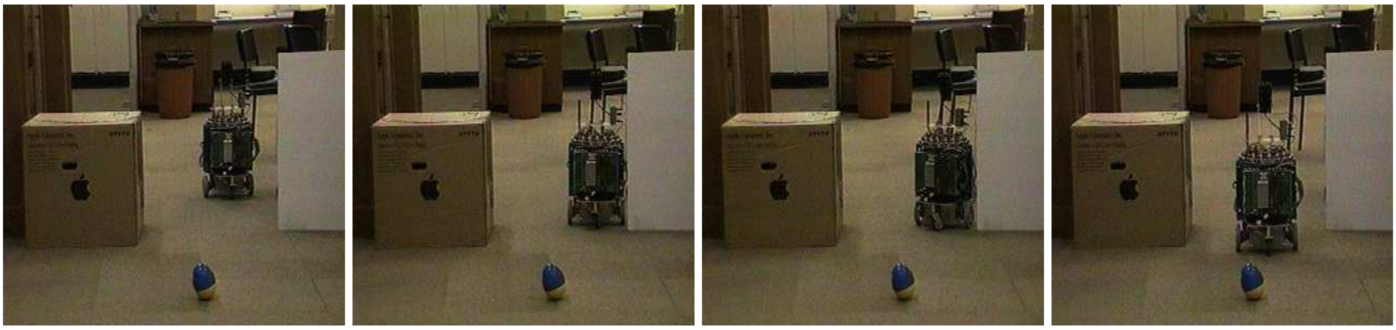

Figure 12. Robot escaping from local minima

Figure 13. Obstacle avoidance

References

[1] I. Horswill, “Visual Collision Avoidance by Segmentation”, Intelligent Robots and Systems, IROS ‘94, pp. 902-909.

[2] T. Gomi and K-I Ide, “Vision Based Navigation for an Office Messager Robot”, Intelligent Robots and Systems, IROS ‘94, pp.2015-2022, Volume 3 1994.

[3] K. Nagatani and S. Yuta, “ Designing a Behaviour to Open a Door and to Pass through a Door-way using a Mobile Robot Equipped with a Manipulator”, Proceedings of Intelligent Robots and Systems Conference, IROS ‘94, pp. 847-853.

[4] Y. Kimura, T. Naito, M.Nakano, H. Moribe and Y. Kuno, “Stereo Vision System for Car Assembly”, Proceedings of IEEE Robotics and Automation Conference, pp 1403-1409, Volume 2 1995.

[5] D. Jung and A. Zelinsky, “Cleaning: An Experiment in Autonomous Robot Cooperation”, The International Symposium on Robotics and Manufacturing, ISRAM, May 1996.

[6] A. Zelinsky and J. Heinzmann, “Real-Time Visual Recognition of Facial Gestures for Human Computer Interaction”, Second International Conference on Automatic Face and Gesture Recognition, October 14-16 Killington , Vermont, USA .

[7] J-C. Latombe, “Robot motion planning”, Boston: Kluwer Academic Publishers, c1991.

[8] H. Gleitman, “Psychology” - Second Edition, 1986, pp 89-90.

[9] P. G. Zimbardo, “Psychology and Life” - Tenth Edition, 1979.

[10] G. Cheng and A. Zelinsky, “A Physically Grounded Search in a Behaviour Based Robot”, Proceedings of the Eighth Australian Joint Conference on Artificial Intelligence. World Scientific, pp. 547-554, 1995.

[11] S. Yuta, S. Suzuki and S. IIda, “Implementation of a Small Size Experimental Self-Contained Autonomous Robot”, Proceedings of the 2nd Int. Symposium on Experimental Robotics, June 25-27, 1991, Toulouse, France.

[12] R.A. Brooks, "A Layered Intelligent Control System for a Mobile Robot", IEEE Trans. on Robotics and Automation, RA-2, April 1986.

[13] O. Khatib, “Real-Time Obstacle Avoidance for Manipulators and Mobile Robots”, International journal of Robotics Research, 5(1), pp 90-98, 1986.

[14] A. Zelinsky, Y. Yuniyoshi, T. Suehiro and H. Tsukune, “Using Augmentable Resource to Robustly and Purposefully Navigate a Robot”, Proceedings of IEEE Robotics and Automation Conference, pp 2586-2592, Volume 3 1995.

[15] S. Iida and S. Yuta, “Vehicle Command System and Trajectory Control for Autonomous Mobile Robots”, IROS ‘91, Nov. 3-5 1991, Osaka, Japan. IEEE Cat. No. 91TH0375-6.

[16] G. F. McCleary and N. Westbook, “Recreational and Re-creational Mapping”, Sturbridge Village, 1974.

[17] H. Schone, “Spatial Orientation: The Spatial Control of Behaviour in Animals and Man”, Princeton, NI: Princeton University Press, 1984.

[18] D. McFarland and T. Bösser, “Intelligent Behaviour in Animals and Robots”, The MIT Press, 1993.

[19] R. Pfeifer, “Cognition - Perspective from autonomous agents”, Robotics and Autonomous Systems 15, pp. 47-70, 1995.

[20] M. Ishii, M. Kakikura and Y. Kuniyoshi, “Mobile Robot Navigation Using Overlapping Environmental Features”, Japan, B204, pp.156-159, 1994.