MVA2002 IAPR Workshop on Machine Vision Applications, Dec. 11 - 13, 2002, Nara- ken New Public Hall, Nara, Japan

3—39 Fast landmark detection system based on the multi-resolution image matching scheme for home robot navigation

Atsushi Shiraishi, Naoyuki Sawasaki
Fujitsu ltd.

Abstract

We devised real-time navigation schemes that enable autonomous travel of a robot on the unoccupied premises of an ordinary house based only on visual information. To accomplish this, we made use of the limitations on the specific operating range of a robot in an empty house. The range can be limited to specific work locations and the paths that connect them. The schemes consist of one with a low degree of accuracy that is used for the travel paths and one with a high degree of accuracy that is used for work locations. The use of low-resolution images helped enable removal of the effects of small changes made daily in the scenes by human activities. The identification of high-resolution characteristic points through application of a coarse-to-fine method ensured the accuracy of self-localization, which is required in work locations. Although such image processing generally requires a large amount of computation, the introduction of hardware dedicated to image correlation provided practical speed.

1: Introduction Now that the average family is becoming smaller, there is an increasing demand for domestic robots designed for assistance in performing household chores. One of the technologies required for domestic robots that we are interested in is the autonomous travel of a robot in an ordinary household while the residents are absent. In this situation, the information that a robot can use for self-navigation is assumed to be only the image information that is received from its cameras. We established this limitation for two reasons.

1) A person, even in an unfamiliar place, can walk around relatively safely without bumping into anything by recognizing the presence of nearby objects and creating a visual image in his or her brain. In a situation where a person is provided with only traveling images through a video camera, he or she can guess the camera location and its travel fairly accurately if the locations shown by the images are known or similar to what is known. In other words, it is clear that there is a method for estimating location using only a memorized map and image information.

2) In terms of reduced cost, a robot should contain as few types and numbers of sensors as possible.

Since image processing such as matching requires a large amount of computation, robot navigation that depends entirely on this method has, so far, been rarely examined. However, this barrier was lowered considerably by "Tracking Vision," specialized hardware for image correlation. We tested a method of matching using this hardware. The following paragraphs discuss the results of the test.

2: Problems We encountered several problems in attempting to have a robot conduct human-like self-navigation based only on visual information. As yet, no scheme is established to automate the creation of a map of the environment based only on visual information in the same way that a human being does so subconsciously. Therefore, a detailed map consisting of coordinate information that is linked with visual information must be input to a robot. This approach requires a large amount of labor and does not lead to a practical solution.

However, it is expected that the tasks that an unattended robot performs at home do not vary greatly from day to day as far as the travel paths, work locations, and work items are concerned. In other words, we can think of the following schemes that do not require detailed map information but only simple instructions and specifications of locations.

- Having a robot memorize only the positions where tasks are performed, including the "home location" and the major "transit points" for traveling (hereafter called "work locations"), and the travel paths connecting these locations as a set of surrounding images seen from the cameras of the robot at these locations and on these paths

This method has the following technical problems:

(a) During autonomous travel, errors must be visually detected so the robot can reach the specified location with sufficient accuracy.

(b) The robot must not be confused by the following two visual characteristics of the human living environment. Furthermore, the use of visual cues designed only to help the robot avoid such confusion does not conform to the premise of using the robot in an ordinary household.

(b-1) Details of the surroundings in an ordinary household are temporally unstable due to daily activities of the residents.

(b-2) In a living space for humans, such as an ordinary household, from one place, a robot might see more than one image that is similar to characteristic points, i.e., possibly effective cues for the robot.

(c) Since image matching generally requires a large amount of computation, a preferable implementation is one that has as little as possible.

3: Proposed Schemes In the proposed schemes, a robot travels basically by measuring the distance to landmarks in images and identifying and correcting its own location. However, we effectively utilize the requirement that different levels of accuracy are employed for self-localization when the robot is on a travel path and when it arrives at a work location. This is because the constant comparison of large numbers of all recorded images and the current surroundings is impractical due to the limitations imposed by the required amounts of computation and storage.

Therefore, we propose the following schemes for travel.

(1) Navigation along a travel path

While traveling along a travel path from one work location to another, a robot records its surroundings as a set of sequenced low-resolution panoramic images. The low-resolution approach avoids problems (b-1) and (b-2), mentioned above, and alleviates problem (c). Problem (a) is not considered here because its solution is only necessary at a "work location."

The navigation procedure used for autonomous travel is

Address: KEN-52, 4-1-1, Kamikodanaka, Nakahara-ku, Kawasaki-city, Kanagawa-ken, 211-8588 Japan.
E-mail: **siraisi@stars.flab.fujitsu.co.jp** **sawasaki@flab.fujitsu.co.jp**

as follows:

- From a database, a robot retrieves a set of low-resolution panoramic images that constitute the path it is to follow.
- During travel, the robot repeatedly performs the following calculations and continuously corrects the direction of travel.
- The robot reduces a front image captured by the camera to create a small, low-resolution rectangular image.
- The robot correlates the reduced image with the low-resolution panoramic images from the database and determines which of the latter has the highest correlation.
- The robot recognizes the vicinity of the location where the chosen low-resolution panoramic image was captured as the current position of the camera (and thus, of the robot).
- In addition, the robot finds the location in the low-resolution panoramic image that best matches the currently captured image and converts the deviation between the location and the center of the low-resolution panoramic image to an angle. The deviation should be close to the deviation between the direction in which the robot is supposed to be looking and in which the camera is actually looking.
- The end of a travel path is the next "work location." When approaching the work location, the robot makes a more accurate "work location arrival decision." The procedure for this is to clip a rectangular area from the center of the image in front of the robot and convert it to the same resolution as the low-resolution panoramic images. This is then reduced or enlarged through several levels to compare it with the past low-resolution panoramic images captured at the "work location." If the reduced front image has a greater match, the robot decides that it has gone past the "work location." If the enlarged front image has a greater match, it has not yet arrived at the "work location." If the front image that is neither reduced nor enlarged has the greatest match, it is approximately at the "work location." Based on these results, the robot fine-tunes its location

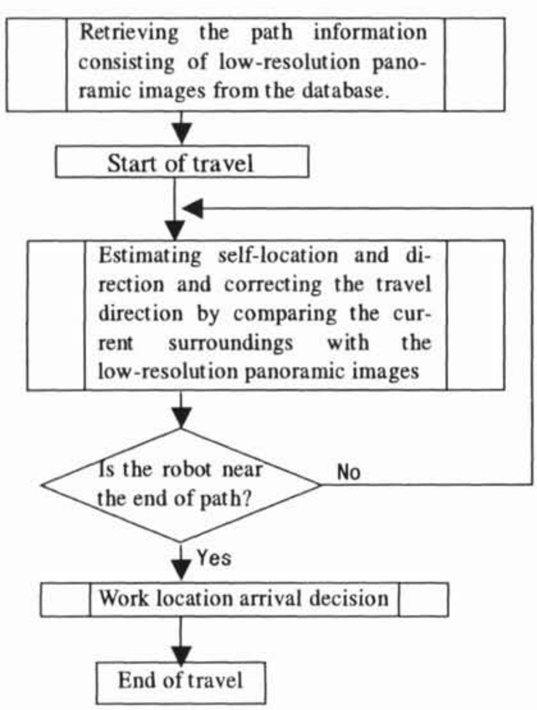


Figure 1 Flow of navigation along a travel path based on low-resolution panoramic images

backward and forward until it is "approximately at the work location." Figure 1 is a flowchart that outlines the procedure of autonomous travel.

(2) **Navigation in a work location**

Details of work locations are stored in a data structure called a "multiresolution feature image map." Figure 2 is an example of a multiresolution feature image map.

A multiresolution feature image map, an application of the coarse-to-fine method, consists of the following three data items:

"Low-resolution panoramic image of surroundings in landscape orientation" --- (1-1)

"Clipped high-resolution images of characteristic points" --- (1-2)

"Numerical data on locations of high-resolution images of characteristic points (1-2) in a low-resolution panoramic image of surroundings (1-1)" --- (1-3)

To teach a work location to a robot, a person first captures a panoramic image of the surroundings as seen from the current location and records it as a low-resolution panoramic image (1-1). Next, the person points to an appropriate image characteristic point around the work location in the robot's view and enters its three-dimensional coordinates in the indoor coordinate system. The robot then clips the image characteristic point from the image in its view and stores it as image information (1-2). The robot memorizes the characteristic point while linking the coordinates in the low-resolution panoramic image (1-1) and the taught three-dimensional coordinates as numerical data (1-3) with the image information of the characteristic point.

The robot, arriving near the work location, retrieves a multiresolution feature image map corresponding to the work location from the database and performs more accurate self-localization as follows:

- The robot searches for as many characteristic points as possible in the surroundings. To do so, the robot:

(2-1) Reduces the resolution of the image in the current view and compares it with low-resolution panoramic images (1-1) through normalized correlation matching to obtain the approximate direction the robot is facing. The use of a low-resolution image in the matching calculation

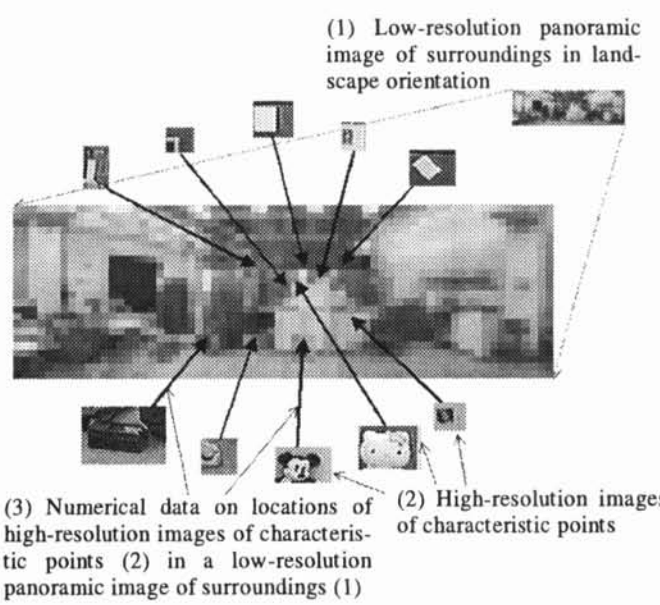


Figure 2 Example of multiresolution feature image map

avoids problem (b-1), mentioned above.

(2-2) Uses the data (1-3) to estimate which characteristic points should be seen in the current view of the robot and their locations.

(2-3) Searches for high-resolution characteristic point images (1-2) that are supposed to be seen in the current view of the robot through normalized correlation matching. During this process, the search for characteristic images should be limited to the range surrounding the points obtained in step (2-2). This limitation avoids problem (b-2), mentioned above.

- After succeeding in rediscovering two or more characteristic points, the robot calculates a more accurate location of itself through trigonometry based on the absolute coordinates and current apparent directions of these characteristic points. This procedure solves problem (a), mentioned above.

The above navigation methods ((1) and (2)) require a large amount of computation because they involve image matching. An ordinary computer cannot easily provide the self-location information to a traveling robot in real time. Therefore, we introduced the use of hardware that is specialized for image correlation operations.

We used the Embedded Tracking Vision "TRV-4," a Fujitsu high-speed image correlation operation system. Tables 1 and 2 list the image operation capabilities and the major specifications of the TRV-4, respectively. Figure 3 is a photograph of the printed circuit board.

4: Experiment We used two environments for experiments. One consisted of a dolly on which a CCD camera, TRV-4, a personal computer, and a monitor were mounted. The other consisted of the robot platform of the Humanoid Robotics Project (HRP)[1] connected to the TRV-4.

(1) Navigation along a travel path

In this experiment, we used the dolly-based environment. As an example of the results of the experiment, Figure 4 shows the "path" consisting of indoor surroundings captured every 47 centimeters during travel along a path and the plane view of how the dolly traveled when the images were captured. The image at the left end is a low-resolution panoramic image captured at the starting point of the path; the image at the right end is a low-resolution panoramic image captured at the ending point of travel, i.e., the goal. These low-resolution panoramic images were captured so that the center of the images is in the desired direction of travel in each of the capture locations.

In the path consisting of nine low-resolution panoramic images, shown in Figure 4, we were able to identify the current location and calculate the current deviation of direction (in degrees) from the desired direction at each point, at a frequency of about three per second. We confirmed that if the dolly is pushed while the traveling direction is continuously adjusted to maintain a zero deviation on the monitor display, the dolly can travel along the instructed path.

In Figure 4, the location where the last and rightmost low-resolution panoramic image was captured was the end of the path, i.e., the next "work location." The robot recognized that this image had the highest correlation with the front image currently captured by the camera, and thus decided that it had approached the next work location, and started to calculate a more accurate arrival decision. We conducted the experiment at the end of path shown in Figure 4. This goal was set at a point 1.5 meters from the wall. Figure 5 shows the images obtained by enlarging and reducing the front image at this location. At this time, the location recognized as the goal was about 5 centimeters forward from the actual location where the low-resolution panoramic image of the goal had been captured.

In other tests, arrival at the goal was generally recognized with an accuracy of 10 centimeters or better. This was true even if the distance to the wall ahead was greater than 10 meters as long as any object closer (within one or two meters) was included in the view.

Table 1 Image operation capabilities of the TRV-4

SAD correlation (8 x 8 pixels)	2390 times per 33 ms
Normalized correlation (8 x 8 pixels)	470 times per 33 ms
Space filter (8 x 8 pixels)	890 times per 33 ms

Table 2 Major specifications of the TRV-4

CPU	ACP Ⅱ
ROM	2MB Flash-EEPROM
RAM	16MB DRAM
External I/O	RS-232C × 2ch
Image input	NTSC-compliant (color), two selectable systems
Image memory	1024×1024×16bit× 8 planes
Power voltage	+5V single
Power consumption	8.5W
Temperature range	0 to 55°C
Dimensions	94mm×90mm

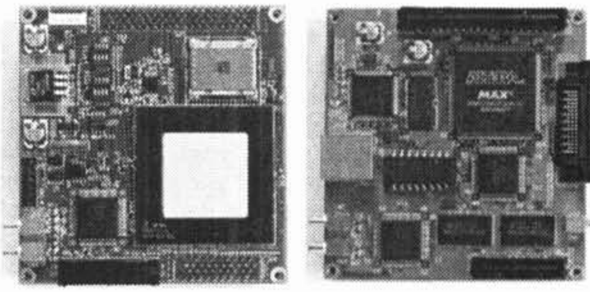


Figure 3 PC board of the Tracking Vision System

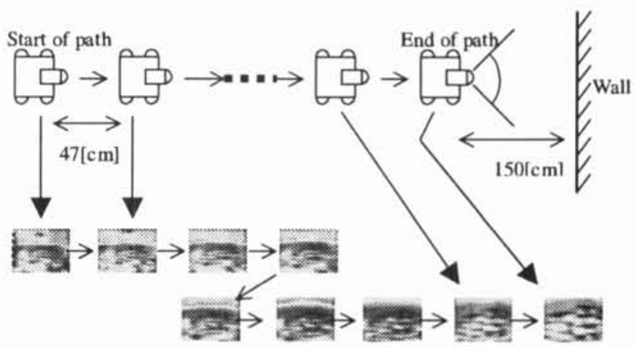


Figure 4 Example of path recorded as a low-resolution panoramic image sequence

Figure 5 Example of enlarged or reduced low-resolution front image used to identify an accurate location near the goal

(2) Self-localization in a work location

We conducted the following experiment using an HRP robot platform with Tracking Vision added.

(2-1) Real-time detection of characteristic points

We were able to detect simultaneously the image areas similar to all 10 high-resolution templates recorded in the multiresolution feature image map, shown in Figure 2. We were also able to confirm that no incorrect detection occurred in a multiresolution feature image map with multiple high-resolution templates that have apparently very similar shapes.

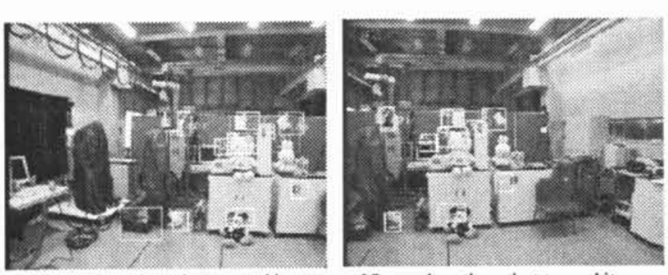
Figure 6 is a photograph of the robot turning its head left to right to vary the line of sight in order to keep track of the characteristic points in the view on a real-time basis. Figure 7 shows how the view looks.

(2-2) Self-localization

Figure 8 shows how the robot estimates self-location in a mockup home space. Figure 9 shows how characteristic points are found in the robot's view. Figure 10 shows the accuracy of self-locations obtained through this method.

The center of Figure 10 indicates the point at which the robot was initially taught the characteristic points of the surroundings. The following degree of self-localization accuracy was obtained through this method.

Figure 6 Robot platform tracing multiple characteristic points using the multiresolution feature image map, shown in Figure 2



View when the robot turned its head halfway to the left — View when the robot turned its head from right to left

Figure 7 Real-time tracing of characteristic points using the multiresolution feature image map during head turning

Figure 8 HRP robot platform estimating self-location

- Location error: Less than 10 centimeters
- Direction error: Less than 2 degrees

(Measured within 50 centimeters from the location at which the characteristic points were taught)

The time required to find the characteristic points, shown in Figure 9, was practically negligible.

5: Conclusion We devised real-time navigation schemes that enable autonomous travel of a robot in an ordinary house while no one was present based only on visual information. To do this, we utilized the limitations on the operating range of a robot in an ordinary, unoccupied household to the specific work locations and paths connecting them. The schemes consist of a low-accuracy one used for paths and a high-accuracy one used for work locations. The use of low-resolution images enabled removal of influences from small changes made in the surroundings by human activities. The identification of high-resolution characteristic points through application of the coarse-to-fine method ensured the accuracy of self-localization that is required in work locations. Although such image processing requires a large amount of computation, the introduction of "Tracking Vision," hardware specialized in image correlation, provided us with a realistic speed.

Acknowledgments This research was conducted as part of the "Humanoid Robotics Project (HRP)," an applied research and development project of the Ministry of Economy, Trade and Industry.

References

[1] H.Inoue et al., HRP: Humanoid Robotics Project of MITI, IEEE Humanoid 2000.

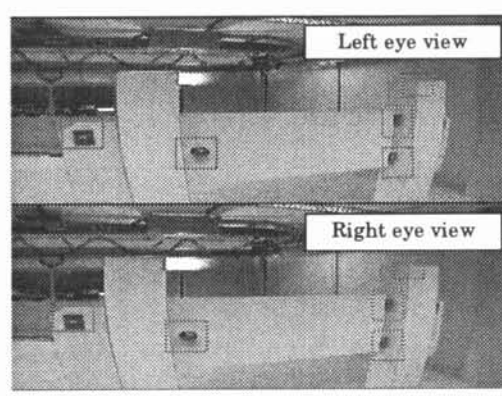


Figure 9 Identifying characteristic points for identifying self-location

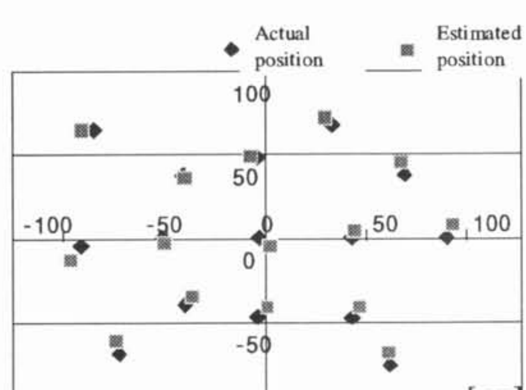


Figure 10 Results of identifying self-locations