Autonomous Navigation of Mobile Robot based on Teaching and Playback Using Trinocular Vision

Akihisa Ohya
University of Tsukuba / PRESTO, JST
1-1-1 Tennodai
Tsukuba 305-8573, JAPAN
ohya@is.tsukuba.ac.jp

Yukio Miyazaki and Shin'ichi Yuta
University of Tsukuba
1-1-1 Tennodai
Tsukuba 305-8573, JAPAN
yuta@esys.tsukuba.ac.jp

Abstract – **The purpose of this study is to realize autonomous navigation of a mobile robot in an indoor environment by teaching and playback method, without making any environmental map in advance. The strategy of navigation in this study is that the robot runs based on its odometry information and corrects its position using landmark information obtained from stereo vision sensor. We use a trinocular vision system to solve easily the stereo matching problem and choose vertical lines in the environment as landmarks because of their easiness to detect and sufficient existence in the environment. In this paper, the method of teaching and playback is described and some experimental results are also shown.**

I. INTRODUCTION

When a mobile robot navigates along a designated path autonomously, in many cases the robot follows the path by recognizing its position in the environment. In the case of a wheeled-type mobile robot, the robot can estimate its position to some extent with the odometry system. But the error of this position estimation is increasing with the motion of the robot. In this case, the robot should adjust its position by using landmarks in the environment. Such navigation methods combining landmark information with an odometry system are very popular subjects in the field of mobile robotics[1-5]. Our group has also done similar research about indoor and outdoor navigation[6,7].

For navigation using landmarks, an environment map that describes positions of landmarks is needed. Such an environment map is commonly made by a human's measurements. It is easy to imagine that to make an environment map requires a lot of time and work. Also, special landmarks, for example laser beacons, should be set up if these are no natural objects.

In a different approach, as reported on by [8,9], a robot is provided with a sequence of images of the interior space. By comparing these pre-recorded images with the camera images taken during navigation, the robot is able to determine its location. In those methods, the information about the designated path is obtained in the form of a sequence of images. The robot can navigate itself without any environment map or any specially prepared landmarks. This is a common advantage of the navigation method having teaching and playback style.

In this paper, we propose an autonomous navigation method for a mobile robot in an indoor environment by a teaching and playback method. In our method, we use only the position of vertical lines in 2D coordinate system extracted by a stereo vision system as landmark information. We use a trinocular vision in order to solve the stereo matching problem simply and quickly. First, the landmark information is extracted while a human controls the robot and it is recorded in teaching stage. While autonomously navigating, the robot corrects the position error by comparing the observed landmark position and the recorded one. This method has the advantage that the robot can run not only on the same path as the taught one but also on a different path because the recorded landmark information is just the position of vertical lines in 2D space and can be used as a general environment map. The robot is able to navigate on an arbitrary path as long as the path is not far from the one indicated during the teaching stage.

In the rest of this paper, we explain in section 2 the strategy of autonomous navigation. The trinocular vision system is described in section 3. The method for teaching and playback presented in section 4 and the experimental results obtained are presented in section 5. Finally the conclusion is given in section 6.

II. STRATEGY OF NAVIGATION

The basic strategy of this navigation method is "teaching and playback." First, as a teaching stage, a human controls a robot by radio or other interface and teaches the path to the robot. Next, as a playback stage, the robot runs autonomously on the path instructed during the teaching stage. During the teaching, the robot records its trajectory and collects information about the position of landmarks using a vision sensor. During the playback, the robot runs autonomously according to the trajectory recorded by the robot and corrects its position comparing the stored landmark information and the observed one.

In this research, we use a trinocular vision sensor, which is an improved stereo vision sensor using three redundant cameras. It can detect landmark and calculate its relative position from the robot. In this study, we use vertical lines on pillars or walls in the environment as shown in Fig. 1. A

Fig. 1. The vertical lines used as landmarks.

vertical line has good characteristics as a landmark. It exists in a natural indoor corridor environment, which is the target environment of this study and it can be expected to be detected at a necessary and sufficient frequency. Also, it is easy to be detected by the stereo vision sensor.

III. TRINOCULAR VISION SENSOR

In this research, we developed a trinocular vision system, which is a kind of stereo vision using three cameras, and used it as a sensor for mobile robot navigation. In usual stereo vision that uses two cameras, it can measure the position of the object by triangulation. It needs to find the correspondence between two images based on the calculation of the correlation between two images. However, because the computation cost of correlation calculation is rather high, this process usually takes a long time. On the other hand, in trinocular vision, the stereo matching can be done in short time using a redundant camera[10]. It is a great advantage for the mobile robot navigation to process images quickly because it requires real time processing.

A. Optical arrangement of the trinocular vision

Fig. 2 shows the optical arrangement of the trinocular vision. Three cameras are set a constant distance apart from each other and their optical axes are parallel to each other in the X-Y plane. The optical axis of the center camera is on the X-axis. In each camera, the roll angle around its optical axis is zero. Since the vertical lines in the environment should appear as vertical edges on the images in such an optical system, vertical edges are easy to be extracted and can be good landmarks. In each image, U-axis is defined as a horizontal axis through the optical center. The edge extraction is performed by finding the points where vertical edges cross the U-axis.

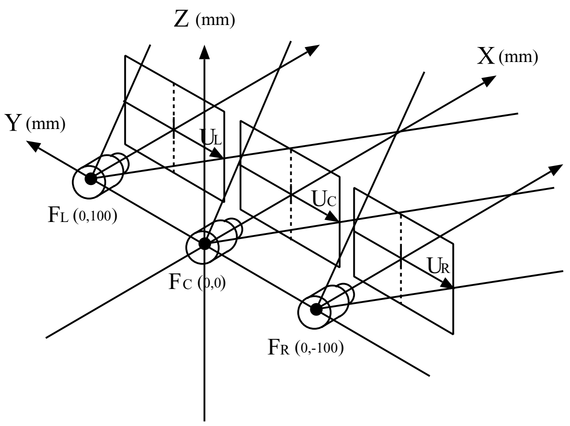

Fig. 2. The optical arrangement of the trinocular vision.

B. Landmark detection

We developed a software to observe landmarks with the trinocular vision sensor. This software finds a suitable vertical line as a landmark and calculates its relative position from the robot. Three camera images from the trinocular vision are captured at the same time. An example of the images is shown in Fig. 3. In the captured images, the landmarks (vertical lines) come out as vertical edges in the images. They are detected by differentiating on the U-axis. The sign of the differentiated value is used in the stereo matching process for robust edge detection. The position of the detected edge is corrected using a lookup table to eliminate the influence of the image distortion.

Fig. 3. An example of the captured images.
(R, C, L express right, center and left, respectively.)

C. Stereo matching in the trinocular vision

The detected edges must be matched between the stereo images. In the trinocular vision, the stereo matching is realized by checking the correspondence between the 1st (R) and the 2nd (L) images referring to the 3rd (C) image instead of carrying out a correlation calculation between the images. As shown in Fig. 4, the position of edge is estimated on the actual edge position in the 2 dimensional space in case that the correspondence in the left and the right camera images is correct so that an edge exists on the expected point in the center camera image. If the correspondence is not correct as shown in Fig. 5, the position of the edge is estimated falsely so that no edge is appearing on the expected point in the center camera image.

The processing procedure is shown below.

1. A correspondence in the 1st and 2nd images is as-

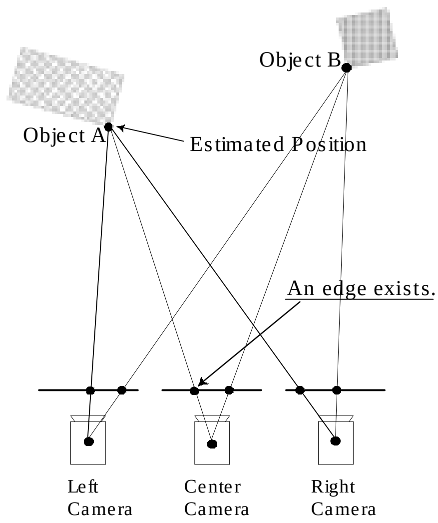

Fig. 4. The correct matching in the trinocular vision.

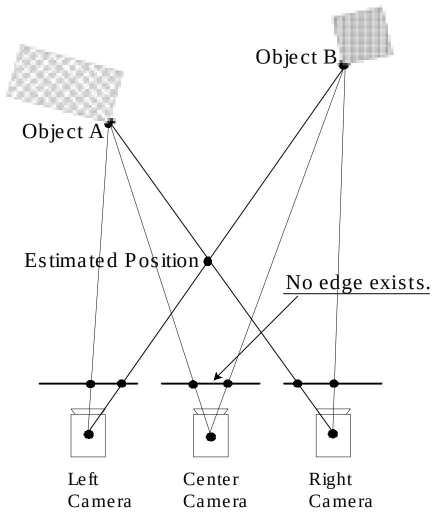

Fig. 5. The false matching in the trinocular vision.

sumed.

2. The position of the landmark is calculated by triangulation based on the assumed correspondence.
3. The estimated edge position in the 3rd image is obtained from the calculated landmark position based on the optical arrangement.
4. If the assumed correspondence is correct, the real edge should exist on the estimated position. Otherwise, we can understand that the correspondence is not correct when no edge exists on the estimated position.

By repeating this procedure, all correct correspondences can be extracted. In this operation, the geometrical restriction of the optical system is also used to shorten the processing time. Also, the consistency of the sign of the differentiated values is checked to avoid mismatching.

D. Usability check of landmark

With the processing described here, the relative position of a vertical line can be obtained as landmark information. However, mismatching still can exist even if the above checks are applied. This will often happen when many vertical lines exist in the environment. Here we can consider that the mismatched edge should match also with another correct edge. It means the mismatched edge must have more than two correspondences. We just throw away such an edge in case multiple correspondences are found because it is difficult to know which correspondence is correct. The purpose of landmark observation is to extract good landmarks for position correction. We don't have to find all of the correct correspondences and only a few landmarks are enough at one observation. Wrong position correction must be avoided, which may happen by mismatching.

IV. NAVIGATION BY TEACHING AND PLAYBACK

A. Teaching stage

The purpose of teaching is to create an environmental map and define a path for the robot before the autonomous navigation. The robot should collect the following two data during the teaching stage.

1. The trajectory of the robot sampled at a moderate distance interval.
2. The position of the detected landmarks in the world coordinate system and the position where the robot observed the landmarks.

The first data is the information concerning the path. It indicates the course where the robot should run during the playback stage. The second information is about the environment. The later half in the second data shows where the robot can observe the landmark during the navigation.

How to collect the above mentioned data in this method is shown below.

1. The robot is moved on the desired path by a human operation (pushing by hand or remote control).
2. The robot records the trajectory using a dead reckoning system (odometry).
3. The positions of the landmarks (the vertical edges in the environment) observed with the trinocular vision sensor are recorded. The positions of the robot at the observation point are also recorded.

The above recordings are performed at individually moderate distance intervals. The distance is determined by each purpose. The robot's position is recorded every 5cm of robot's movement. The landmark observation is attempted when the robot runs 1m from the previous position where the landmark was detected and recorded. If no landmark is found by the observation, the robot tries again to detect the landmark until at least one landmark is found. Since this recording process will be done automatically on the robot, the human operator who teaches the path doesn't have to care about the sensing status.

In the case where more than one vertical line was found out by a landmark observation and these lines are close enough to each other, we have decided to throw away all of them in order to secure the landmark matching between teaching and playback. See later for details.

B. Playback stage

In the playback stage, the robot runs autonomously according to the taught data. The tasks to do during this stage are shown below.

1. The robot tries to move on the same path as the teaching stage according to the recorded data.
2. The robot tries to observe the landmark when the robot approaches the place where the robot detected the landmark. (Fig. 6 shows the condition. The observation is started when the robot enters the gray area.)
3. The robot corrects its position by comparing the observed landmark position and the recorded position.

For the first task, the robot moves relying on the odometry. However, it is impossible to succeed in the navigation because the position error is influenced significantly by the initial error of the robot's orientation, and the error of odometry also will be accumulated gradually. This problem can be overcome by the second and the third tasks. The error will be cancelled during the navigation using landmark information.

We describe about how to correct a position concretely when it is possible to observe a landmark. At first, the robot should specify which landmark is detected in the recorded landmark list when it could find a landmark. This problem can be solved by simply checking the distance between the

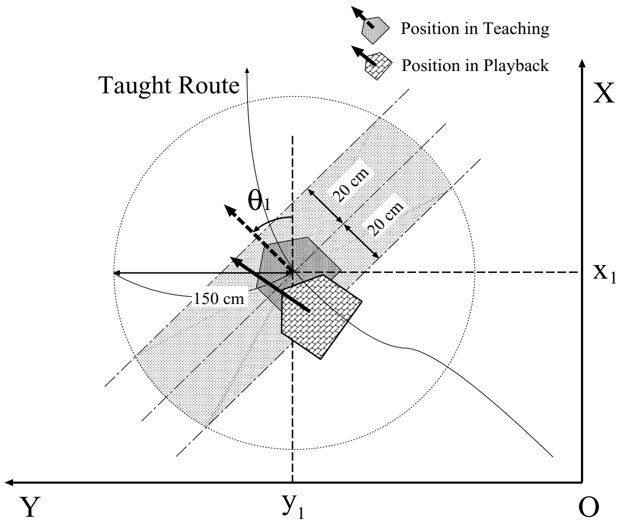

Fig. 6. The condition for the landmark observation in the playback navigation.

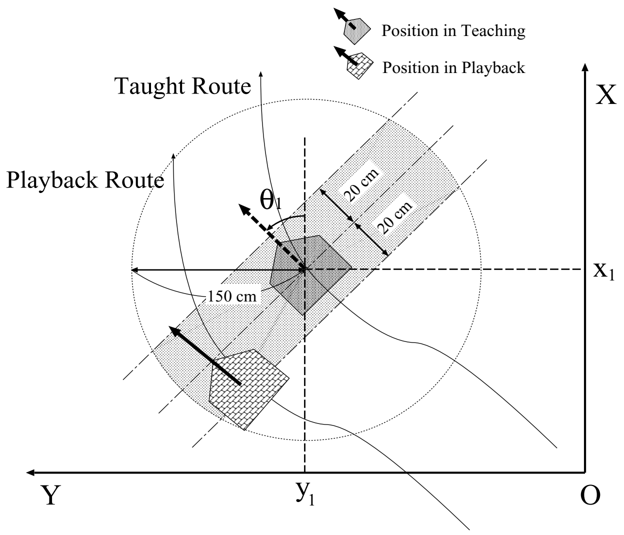

Fig. 7. The condition for the landmark observation during the playback with an offset.

detected landmark position and the stored position. This will be valid assuming that the position error of the robot is remaining small by the position collection at a moderate interval. To ensure this, when more than one landmark are found in a small area during the landmark detection process, they should be thrown away both during the teaching and the playback stage. Generally, the actual position of the robot during the playback stage is slightly different from the position in teaching stage even though the estimated position is the same as teaching. In this case, the position of the observed landmark must be different. This difference comes from the difference of the robot's position. Therefore the error can be corrected using this difference information. The position is managed in the robot by a position estimation system developed in our laboratory[11]. The position is estimated in this system in association with the covariance matrices and it will be corrected using landmark observation information based on Maximum Likelihood Estimation method.

C. Extension of the playback navigation

In the previous section, the playback is described as the navigation on the same route as the one of the teaching. However, the robot doesn't need to follow exactly the same trajectory. Theoretically this method can be valid for the navigation along an arbitrary path as long as the robot can correctly observe the landmarks that are recorded during the teaching stage. It means that the robot should move on the route near the taught trajectory. As examples, the following extensions of the method are considered.

1. Navigation on a path which has an offset coordinate value. (Fig. 7 shows the condition for the landmark observation during the playback with an offset.)
2. Navigation on an arbitrary path indicated manually near the taught trajectory.

V. NAVIGATION EXPERIMENTS

We made several experiments with an autonomous mobile robot using the proposed navigation method. As an autonomous mobile robot, we use the "Yamabico" robot that we have designed and developed[12]. The Yamabico has a vision and a locomotion control system. The vision system can grab images of 792 x 240 pixels from three image sources simultaneously. The trinocular vision sensor is mounted on the top of the robot as shown in Fig. 8. Three CCD cameras are set 10cm apart from each other and work synchronously. The field of view is about 80 degrees in the horizontal axis.

A. Teaching and Playback

The experimental environment is an indoor corridor shown in Fig. 9. The taught path was approximately 10m long including zigzag parts. The experimental result is shown in Fig. 10. The trajectories for the teaching and playback are presented together with the shapes of the surrounding objects. Fig. 11 is a magnified version in Y direction for better understanding. It shows only the trajectories. The discontinued parts of the recorded path of playback navigation show the places where the position correction was performed. The robot could navigate autonomously referring to the landmark information which was

Fig. 8. The trinocular vision sensor mounted on the robot.

stored during the teaching stage. The errors at the final stop position were -5.1cm in X direction and +3.3cm in Y direction.

Fig. 9. The experimental environment.

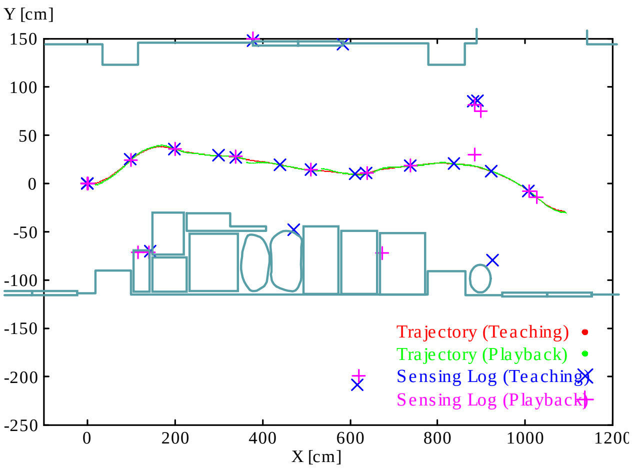

Fig. 10. The trajectories for the teaching and playback experiment.

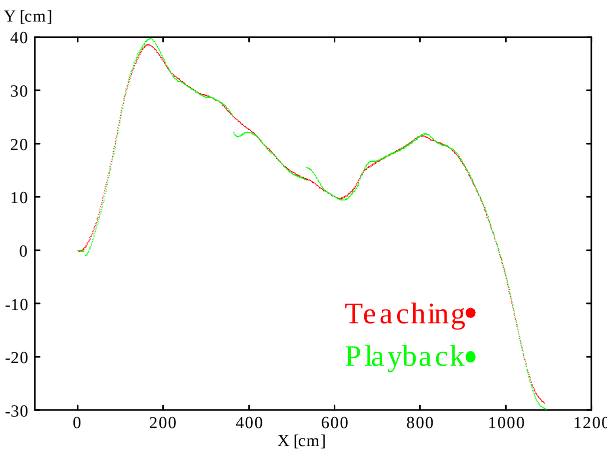

Fig. 11. The trajectories for the teaching and playback experiments. (It is magnified only in Y direction.)

B. Playback with an Offset

The experiment of playback with an offset was performed in the same corridor as the previous experiment. The taught path was a 30m round trip course and the robot ran on the trajectory with a constant offset to the taught path. The offset was 0cm in X direction and +30cm in Y direction. The data of the teaching stage and the playback stage are shown in Fig. 12 and Fig. 13, respectively. The data contains the trajectory and observed landmark information. To compare them, those data are displayed together in Fig. 14.

It can be seen that the actual robot's trajectory for the playback with an offset is parallel to the one for the teaching and it keeps a constant offset given to the robot in the playback stage. The differences at the final stop position from the taught position were -3.5cm in X direction and +27cm in Y direction. This means that the error were -3.5cm in X and -3cm in Y. By considering the total path length, it is possible to say that the system has enough precision.

VI. CONCLUSION

We proposed a navigation method for an autonomous mobile robot in an indoor environment based on a teaching and playback style in which a stereo vision sensor is used for obtaining landmark information. We use a trinocular vision to shorten the processing time for the stereo matching. We implemented the navigation method on our mobile robot equipped with the trinocular vision system and proved the effectiveness of the proposed method through experiments in an indoor corridor environment. The proposed navigation method doesn't require a precise environment map built by human in advance.

We proposed as well an extended navigation method in which the robot can run on a different path from the one of the teaching stage. We proved the applicability of the method by presenting an experimental result of the playback navigation with an offset value in 2D coordinate system.

ACKNOWLEDGMENT

The authors would like to thank Mr. Launay Fabien of the University of Tsukuba for his useful suggestions for preparing this manuscript.

REFERENCES

[1] K. Hyppä, "Optical navigation system using passive identical beacons," in *Proceedings of the Intelligent Autonomous Systems*, 1986, pp.737-742.

[2] K. Sugihara, "Some location problems for robot navigation using a single camera," *Computer Vision, Graphics and Image Processing*, vol. 42, no. 1, Apr. 1988, pp. 112-129.

[3] J. J. Leonard and H. F. Durrant-Whyte, *Directed sonar sensing for mobile robot navigation*, Kluwer Academic Publishers, 1992.

[4] A. Kosaka and A. Kak, "Fast vision-guided mobile

robot navigation using model-based reasoning and prediction of uncertainties," *CVGIP--Image Understanding*, vol. 56, no. 3, Nov. 1992, pp. 271-329.

[5] A. J. Muñoz and J. González, "Localizing mobile robots with a single camera in structured environments," in *Proceedings of World Automation Congress -Robotic and Manufacturing Systems*, May. 1996, pp.539-544.

[6] T. Tsubouchi and S. Yuta, "Map assisted vision system of mobile robots for reckoning in a building environment," in *Proceedings 1987 IEEE International Conference on Robotics and Automation*, pp. 1978-1984.

[7] S. Maeyama, A. Ohya and S. Yuta, "Positioning by Tree Detection Sensor and Dead Reckoning for Outdoor Navigation of a Mobile Robot," in *Proceedings of the 1994 IEEE International Conference on Multisensor Fusion and Integration for Intelligent Systems*, pp.653-660.

[8] Y. Matsumoto, M. Inaba and H. Inoue, "Visual navigation using view-sequenced route representation," in *Proceedings 1996 IEEE International Conference on Robotics and Automation*, pp. 83-88.

[9] T. Ohno, A. Ohya and S. Yuta, "Autonomous navigation for mobile robots referring pre-recorded image sequence," in *Proceedings of the 1996 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 672-679.

[10] Y. Kitamura and M. Yachida, "Three-dimensional data acquisition by trinocular vision," in *Advanced Robotics*, vol. 4, no. 1, 1990, pp. 29-42.

[11] S. Maeyama, A. Ohya and S. Yuta, "Non-stop outdoor navigation of a mobile robot --retroactive positioning data fusion with a time consuming sensor system--," in *Proceedings of the 1995 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 130-135.

[12] S. Yuta, S. Suzuki and S. Iida, "Implementation of a small size experimental self-contained autonomous robot -sensors, vehicle control, and description of sensor based on behavior-," R. Chatila et al. Eds, *Experimental Robotics II*, Springer-Verlag, 1993, pp. 344-358.

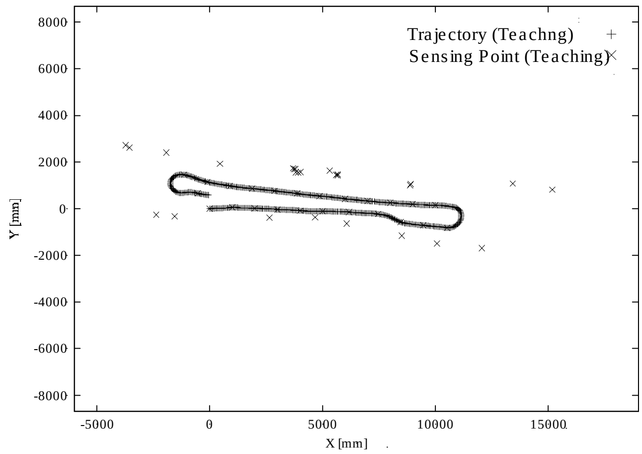

Fig. 12. The trajectory and the observed landmarks information during the teaching stage.

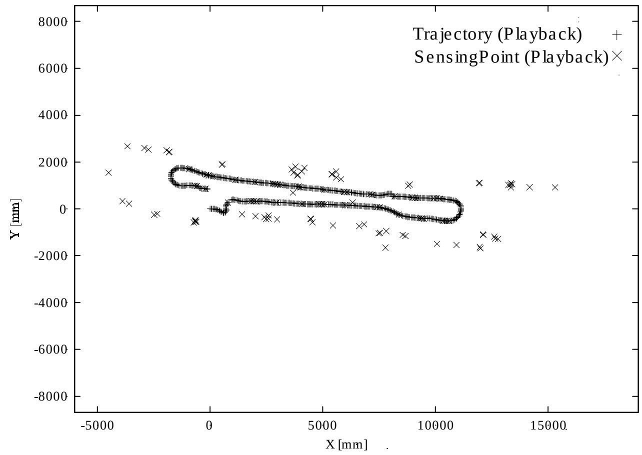

Fig. 13. The trajectory and the observed landmarks information during the playback with an offset.

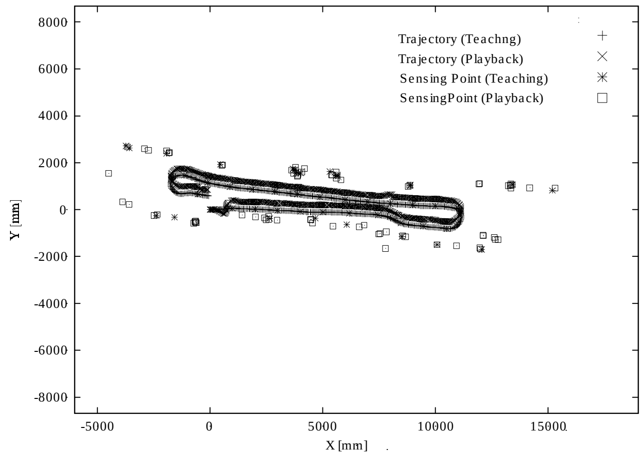

Fig. 14. Trajectory and observed landmarks information for both the teaching stage and the playback with an offset.