I-012

LOCALIZATION OF HUMANOID ROBOT FROM MONOCULAR IMAGE SEQUENCE USING TWO LARGE TEMPLATE MATCHING

Ni Ni Soe† Kenji Suzuki‡ Shuji Hashimoto‡ Jun Ohya†

†Graduate School of Global Information and Telecommunication Studies, Waseda University
‡ Dept. of Applied Physics, Waseda University, Tokyo, Japan

Abstract

In the mobile robot navigation, computer vision technique has been applied to obtain accurate position estimation. In this paper we propose a monocular camera based algorithm to estimate the absolute position of the robot from the image sequences, given initial position and environmental dimension surrounding of the robot. The absolute position is then estimated from geometrical calculation and integration of matching correspondence and environmental data. We did some experiments with a human-scale humanoid robot and verified our method could aid to the navigation of the robot.

1. Introduction

In order for a mobile robot to perform its assigned tasks, it often requires a representation of its environment, a knowledge of how to navigate in its environment, and a method for determining its position in the environment. This paper is concerned with a question of "Where am I?", that of position estimation, which is commonly referred to as *localisation*. Localization is the process by which a mobile robot or other physical agent keeps track of the position as it moves around an environment. Knowing the position and orientation in the operating environment is an essential capability for any robot exhibiting coherent, goal-oriented behavior. The self-localization problem has been studied intensively in the mobile robot community, and a wide variety of approaches with distinct methods and capabilities have been developed [eg.3].

2. Approach

This paper presents new algorithm of absolute position estimation of mobile robot from a monocular image sequence using two feature areas in each image, while the robot is acting in the environment. Mainly, there are three steps in position estimation algorithm: 1) Finding displacement vector in consecutive frames, we use two areas illustrated in the left figure of Fig 2. 2) Finding change in position by the geometrical data (l) and initial pose (x_0, y_0). 3) Integrating into current position ($x_0+\Delta x$, $y_0+\Delta y$, $\theta_0+\Delta\theta$). The procedure is illustrated in Fig. 1.

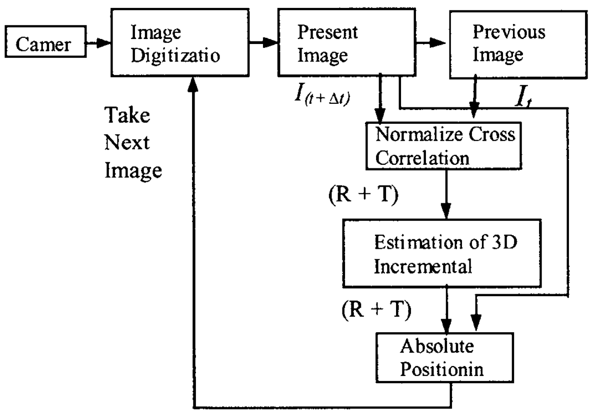

Fig. 1 Data flow for position estimation algorithm

3. Optical Flow Estimation

An optical flow algorithm estimates the 2D flow field from image intensities. We used a region based matching techniques [2] to estimate the optical flow, using normalized cross correlation for efficiency reasons. Consider an image of size $I\times I$, a patch radius of P, and search range of R. We define a displacement vector D on the images I_1 and I_2 for the pixel (x, y) and disparity (dx, dy) as:

$$D(x, y, dx, dy) = \sum_{i=-P}^{P}\sum_{j=-P}^{P} E$$

where,

$$E = \frac{[I_1(x_1+i, y_1+j) - \overline{I_1(x_1, y_1)}]\times[I_2(x_2+i, y_2+j) - \overline{I_2(x_2, y_2)}]}{(2p+1)(2p+1)\sqrt{\sigma^2(I_1)\times\sigma^2(I_2)}}$$

$$\sigma(I_t) = \sqrt{\frac{\sum_{i=-P}^{P}\sum_{j=-P}^{P} I_t^2(x, y)}{(2P+1)(2P+1)} - \overline{I_t(x, y)}}$$

To determine the flow between images I_1 and I_2 at a point (x, y) we maximize $D(x, y, dx, dy)$ for $(dx, dy)\in[-R\times R]\times[-R\times R]$.

4. Coordinate System of the Environment

In our algorithm, we use the following model of the environmental structure as shown in Fig.2. The robot generates the current position from a prior knowledge of environment such as initial position (x_0, y_0, θ_0), and the lateral distance of the structure of the environment (l).

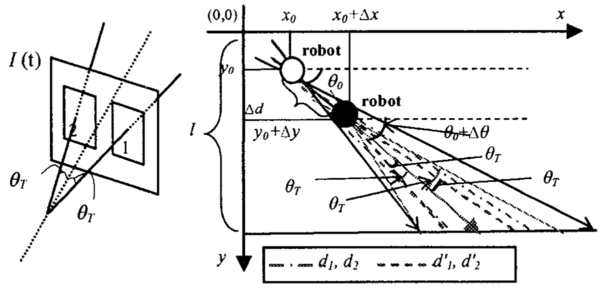

Fig.2 The coordinate system of the surrounding environment

5. Data Association

After computing the displacement vector in two template size, we integrate these values to find the correspondence position of each step of robot movement. If $(x+\Delta x, y+\Delta y, \theta+\Delta\theta)$ is the new location of the feature at time $t+\Delta t$, then the trajectory of motion is given by:

$I(x, y, \theta, t) = I(x+\Delta x, y+\Delta y, \theta+\Delta\theta, t+\Delta t)$

1) Computing the orientation difference $\Delta\theta$ in each step

$$\Delta\theta = \tan^{-1}\left[\frac{\sin(\theta_0+\theta_T)\sin(\theta_0-\theta_T)[(D_2+d_2)-(D_1+d_1)]}{(D_1+d_1)\sin(\theta_0-\theta_T)\cos(\theta_0+\theta_T)-(D_2+d_2)\sin(\theta_0+\theta_T)\cos(\theta_0-\theta_T)}\right]$$

where D_1 and D_2 are the displacement vector of each template size; d_1 and d_2 are the size of feature on the image plane; θ_0 is the initial orientation and θ_T is the half of view angle.

2) Estimating the Translational Position Difference Δx, Δy, Δd

$$\Delta y_1 = l - y_0 - \left[\frac{(l - y_0)d_1 \sin(\theta_0 + \Delta\theta - \theta_T)}{(d_1 + D_1)\sin(\theta_0 - \theta_T)} \right]$$

$$\Delta y_2 = l - y_0 - \left[\frac{(l - y_0)d_2 \sin(\theta_0 + \Delta\theta + \theta_T)}{(d_2 + D_2)\sin(\theta_0 + \theta_T)} \right]$$

where Δy_1 and Δy_2 are changes in point in y- axis; y_0 is initial pose in y-axis; l is the lateral distance from the initial pose to the scene. From these two position changes, we can drive the other changes in point in x- axis and change in displacement;

$$\Delta y = \Delta d \sin \theta_0$$

$$\Delta x = \Delta d \cos \theta_0$$

3) Robot's current relative orientation and position (x_i, y_i, θ_i) at each time interval by integrating with the initial position of the robot;

$$\theta_i = \theta_0 + \Delta\theta$$

$$x_i = x_0 + \Delta x$$

$$y_i = y_0 + \Delta y$$

6. Experimental Setup and Results

The feature area based algorithms have features that effectively cover a greater proportion of the image plane and therefore are more accurate. Several experiments have been performed in order to validate the feature-based matching strategy and estimate the accuracy of estimation by comparing them with the ground truth (motion capture system). Humanoid robot situations have been used to test our system as shown in Fig 3. The images are acquired by a color camera carried by the iSHA humanoid robot [1]. The image acquisition frame rate is set to 10 f.p.s. All the experiments were carried out in the laboratory and the robot was human driven. The robot was positioned in the indoor environment about 380 or 440 cm lateral distance to the target scene, initial orientation is about 92° or 30°, and half of view angle θ_T is 18.5°.

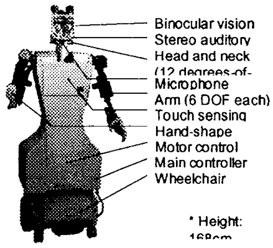

Fig.3 Autonomous humanoid robot "iSHA"

6.1. Pure Translational Motion

In this experiment, the robot moves straight forward in a straight line, with very little rotational velocity. The results are shown in Fig .4 and Fig.5 for translation and rotation respectively. The accuracy is shown in Table 1.

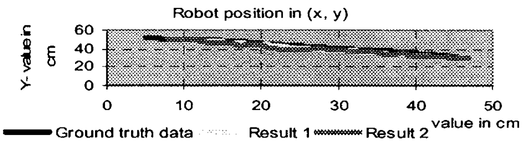

Fig. 4 Position estimates in translation and corresponding ground truth for x and y position change image sequence. Result 1: Template size 60 ×110, Searching range : 25, Frame interval : 1 Result 2 : Template size 60 × 100, Searching range : 30, Frame interval :4.

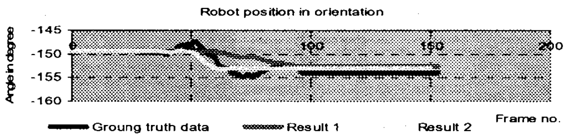

Fig.5 Position estimates in orientation and corresponding ground truth for x and y position change image sequence.

Experiment	Result 1	Result2
Real Displacement(cm)	120	120
Measurement Displacement(cm)	113.238	111.987
Error (%)	5.635	6.68

Table 1 Accuracy for position estimation for forward movement

6.2. Pure Rotational Motion

The motion occurring in the rotation test is much more difficult to estimate. The test area was the same as for the straight movement experiment. The result of various template sizes is shown in Fig. 6 and accuracy is shown in Table 2.

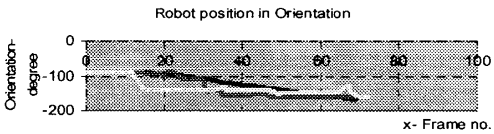

Fig. 6 Position estimates in rotation and corresponding ground truth for small angle rotation sequence. Result 1: Template size 40 ×40, Searching range : 90, Frame interval : 1. Result 2 : Template size 60 × 70, Searching range : 100, Frame interval :1.

Experiment	Result 1	Result 2
Real movement (degree)	76.357	76.357
Measurement angle (degree)	68.8	68.1
Error (%)	9.89	10.8

Table 2 Accuracy for position estimation of initial rotation angle -92.5 degree to final orientation -168.857 degree

7. Conclusion and Future Works

This method uses an imaging technique with a single low-cost CCD camera along with a reference environment to calculate estimation for relative 3D position and orientation of the robot. Experimental results on real environments indicate that the improvements significantly increase robustness precision of self-localization and very promising for practical, real-world implementation. The result is improved state estimates with less error. This method gives good results on localizing of a mobile robot especially in pure translation. Hopefully this method can be implemented so as to work in real time simutaneously for localization and map building.

References

[1] K. Suzuki, et.al., "Development of Autonomous Humanoid Robot iSHA for Harmonized Human Machine Environment," J. Robotics and Mechatronics, Vol.14, No.5, pp. 324-332, 2002.

[2] Barron, J.L., Fleet, D.J. and Beauchemin, S.S., "Performance of Optical Flow Techniques", International Journal of Computer Vision, 12:1, pp. 43-77, 1994.

[3] R. M. Haralick and H. Joo, "2D-3D Pose Estimation", In Proceedings of the International Conference on Pattern Recognition, pp. 385-391, 1988.