Full Length Research Paper

Visual compensation in localization of a robot on a ceiling map

T. Matsumoto[1], T. Takahashi[1], M. Iwahashi[1], T. Kimura[1], S. Salbiah[2] and N. Mokhtar[2]*

[1]Faculty of Engineering, Nagaoka University of Technology, Nagaoka, Niigata, Japan.
[2]Department of Electrical Engineering, Faculty of Engineering, University of Malaya, Kuala Lumpur, Malaysia.

In a visual odometry system, location of a mobile robot is automatically estimated (localized) from video. When the video is captured by an "upward" camera fixed to an indoor mobile robot, a panorama image of the ceiling (ceiling map) is generated by using a visual motion between two adjacent frames in the video. Similarly, location of another robot can be estimated on the ceiling map by using a visual motion between the current frame and the previously generated ceiling map. Under the assumption that the robot goes straight or rotates around a fixed point, there is no problem on the localization as far as the floor is flat. However, when there is debris on the floor, the estimated location contains error. In this paper, we reduce this error by utilizing visual motions in video from the "forward" camera fixed to the robot. This is a visual compensation of motions in the "upward" camera's video with those in the "forward" camera's video. It was experimentally confirmed that the maximum absolute value of the error was reduced to approximately 11%.

Key words: Visual odometry, localization, robot vision, motion estimation.

INTRODUCTION

Auto-localization of a moving robot is a challenging problem for engineers to develop a robot vision network, in which various kinds of mobile robots communicate and cooperate with each other automatically. In a situation such that a global positioning system (GPS) is not applicable, various inner robot sensors such as an acceleration sensor, a geomagnetic attitude sensor, a laser range finder and a gyroscope are useful for this purpose.

Comparing to these sensors, video based localization methods are less accurate and less robust in general. However, it is still attractive for constructing an energy saving small mobile robot with micro CCD sensors. It is becoming useful owing to recently developed odometry and simultaneous localization and mapping (SLAM) techniques (Desouza and Kak, 2002; Nister et al., 2004; Munguia and Grau, 2007).

In this paper, we deal with a system in which location of a mobile robot is automatically estimated (localized) from video. When the video is captured by an "upward" camera fixed to an indoor mobile robot, a panorama image of the ceiling (a ceiling map) is generated by using a visual motion (motion vector) between two adjacent frames in the video. Similarly, location of another robot can be estimated on the ceiling map by using a motion vector between the current frame and the previously generated ceiling map (Papanikolopoulos et al., 1993). Under the assumption that the robot goes straight or rotates around a fixed point, there is no problem on the localization as far as the floor is flat (Udomsiri et al., 2009). However, when there is debris on the floor, the estimated location contains error.

In this paper, we reduce this error by utilizing motion vectors in video from the "forward" camera fixed to the robot. This is a visual compensation of motions in the "upward" camera's video with those in the "forward" camera's video. We theoretically analyze relation between kinetic movements of a robot and motion vectors observed in videos from the upward camera and the forward camera.

To estimate a motion vector, correlation based matching technique has been utilized not only for MPEG video data compression standard, but also for map generation and

---

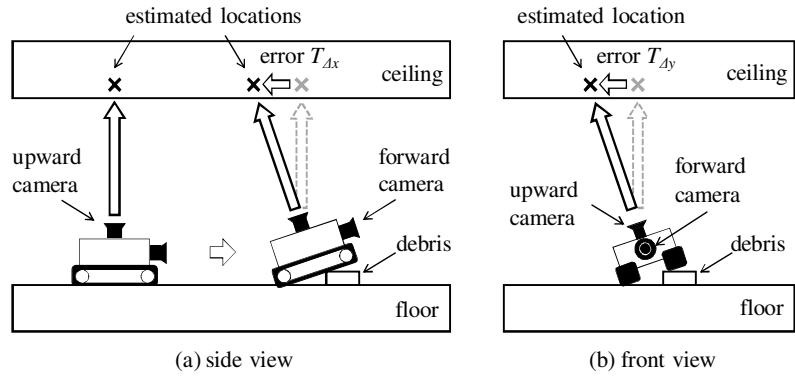*Corresponding author. E-mail: norrimamokhtar@um.edu.my.

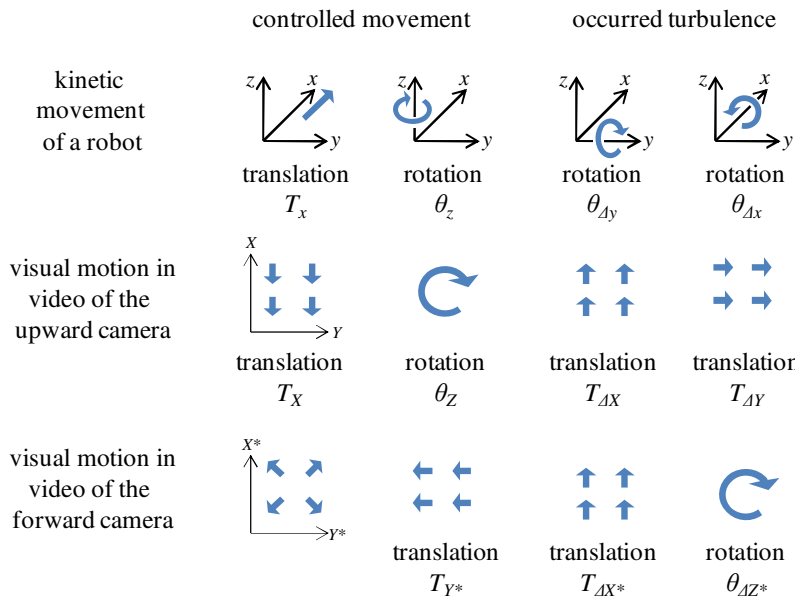**Figure 1.** Localization of a robot: (a) side view, (b) front view.

Figure 2. Movements of a robot and their corresponding visual motions in video.

localization (Wilson and Theriot, 2006). However it is sensitive to lighting conditions. Therefore, we utilize the rotation invariant phase only correlation (RI-POC) technique for stable and precise motion estimation (Sasaki et al., 1998; Ito et al., 2004).

OVERVIEW OF THE SYSTEM AND ITS PROBLEM

Figure 1 illustrates an indoor mobile robot. It is controlled to go straight or rotate around a fixed point on a flat floor under a flat ceiling. The robot has two video cameras - upward and forward. It is our purpose to automatically and precisely estimate its location (localization) from the videos.

Firstly, a panorama image of the ceiling (a ceiling map) is generated by using a visual motion (motion vector) between two adjacent frames in video from the upward camera. Secondly, location of another robot is estimated on the ceiling map by using a motion vector between the current frame and the previously generated ceiling map.

Under the assumption that the robot goes straight or rotates around a fixed point, there is no problem on the localization as far as the floor is flat. However, when there is debris on the floor as illustrated in the figure, the estimated location contains error $T_{\Delta x}$ and/or $T_{\Delta y}$.

ANALYSIS AND THE PROPOSED METHOD

Analysis on kinetic movements and visual motions

Figure 2 summarizes relation between kinetic movements of a robot and visual motions (motion vectors) observed in videos from the two video cameras. Under the assumption in the previous section, a robot is controlled to have only two kinetic movements - translation T_x and rotation θ_z. In this case, in video from the upward camera, translation T_X and rotation θ_Z are observed. It determines the
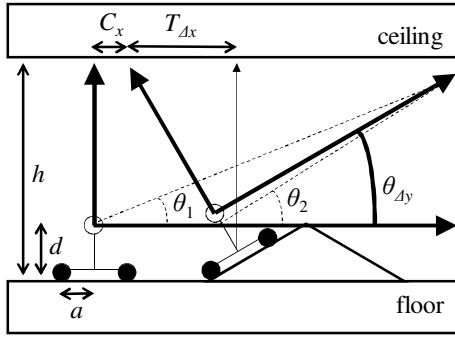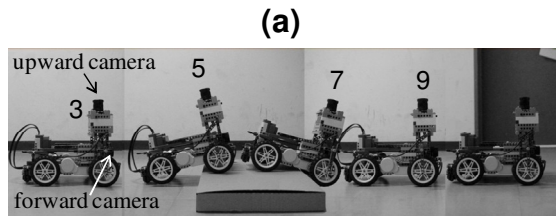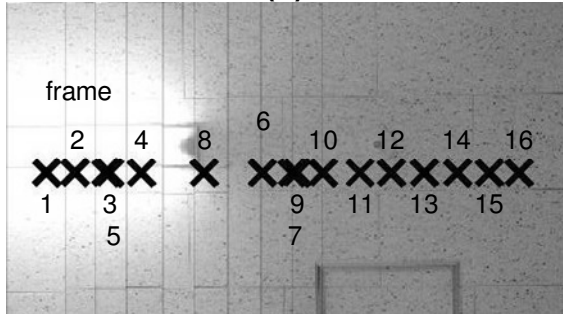
Figure 3. Modeling of translation distances and rotation angles.

(a)

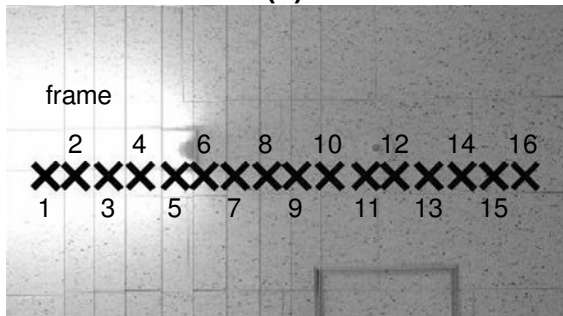**(b)**

(c)

Figure 4. Experiment 1. (a) side view (b) before (c) after.

controlled movements T_x and θ_z without any error.

However, if there is debris on the floor, the robot has two turbulences - rotation $\theta_{\Delta y}$ and rotation $\theta_{\Delta x}$. Each of them generates translation error $T_{\Delta X}$ and $T_{\Delta Y}$ respectively. In this case, mixture of T_X and $T_{\Delta X}$ is observed in X direction in the video from the upward camera. Those are not separable by using only the upward camera.

It is our purpose to compensate this error for precise estimation of movement and location of a mobile robot.

VISUAL COMPENSATION

In this paper, we utilized visual motions in video from the "forward" camera to compensate the error. This is a visual compensation of motions in the "upward" camera's video with those in the "forward" camera's video. We estimate the visual motions ($T_{\Delta X'}$, $\Delta\theta_{Z'}$) in the forward camera's video to calculate the errors ($T_{\Delta x}$, $T_{\Delta y}$) for precise estimation of (T_x, θ_z).

Figure 3 illustrates a model for theoretical analysis on our visual compensation. According to the figure, we have a relation between parameters as

$$d - \left(d + a\tan\theta_{\Delta y}\right)\cos\theta_{\Delta y}$$
$$+ \left\{\left(h + \frac{C_x}{\tan\theta_{\Delta y}}\right)\tan\theta_{\Delta y} - \left(d + a\tan\theta_{\Delta y}\right)\sin\theta_{\Delta y}\right\}$$
$$\times\left\{\tan\theta_{\Delta y} + \frac{\tan\theta_2}{\tan\theta_1 - \tan\theta_2}\left(\tan\theta_{\Delta y} - \tan\theta_1\right)\right\}$$
$$= 0$$

$$(1)$$

where the distances a, d and h are previously measured constant values. This equation determines the error as

$$T_{\Delta x} = (h - a\sin\theta_{\Delta y})\tan\theta_{\Delta y}. \tag{2}$$

Similarly, we have

$$T_{\Delta y} = (h - a\sin\theta_{\Delta x})\tan\theta_{\Delta x}. \tag{3}$$

In the proposed method, we calculate the rotation turbulences ($\theta_{\Delta y}$, $\theta_{\Delta x}$) in the equation above from the visual motions ($T_{\Delta X'}$, $\Delta\theta_{Z'}$) estimated by using video from the forward camera. The estimation is carried out by the rotation invariant phase only correlation (RI-POC) described in [7,8]. As a result of utilization of the forward camera's video, it is expected to have reduced errors in localization of a robot on the ceiling map.

EXPERIMENTAL RESULTS

In the following experiments, the parameters in Figure 3 are a = 14 cm, d = 45 cm, and h = 260 cm, respectively. The robot goes straight at a constant speed.

Example 1

Figure 4a illustrates side view of this experiment. The debris brings about the error $T_{\Delta x}$ to be eliminated by the proposed method. Figure 4b and 4c illustrate the estimated locations on a ceiling map before and after applying the proposed method respectively. In Figure 4c, locations are plotted in a line at a constant interval as the errors are eliminated and the locations are compensated.

Figure 5 summarizes the error $T_{\Delta x}$ for the case in Figure 4. It is observed that the maximum absolute value of the
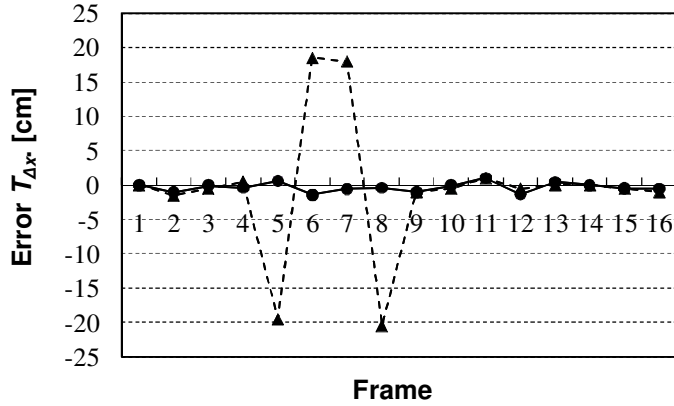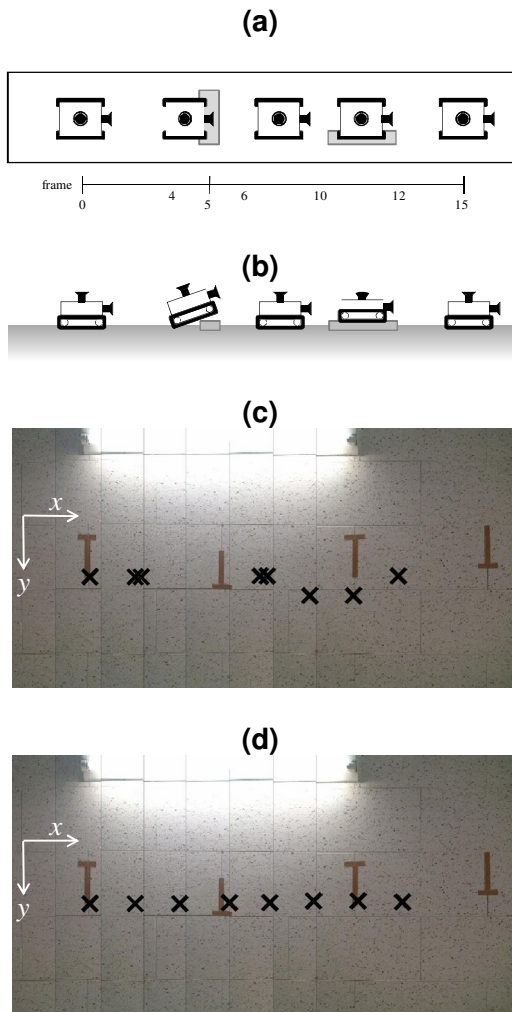
Figure 5. Error in the experiment 1.

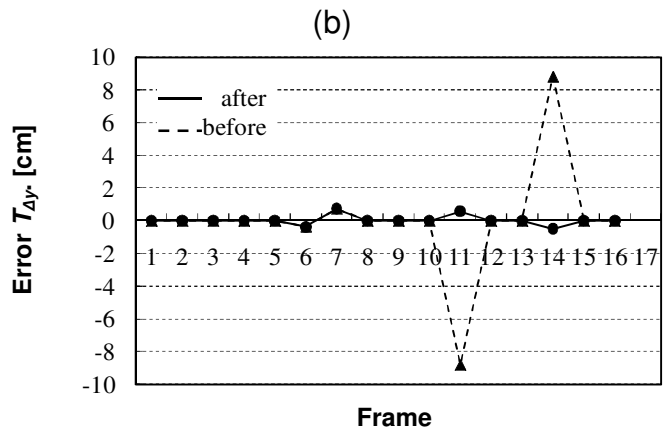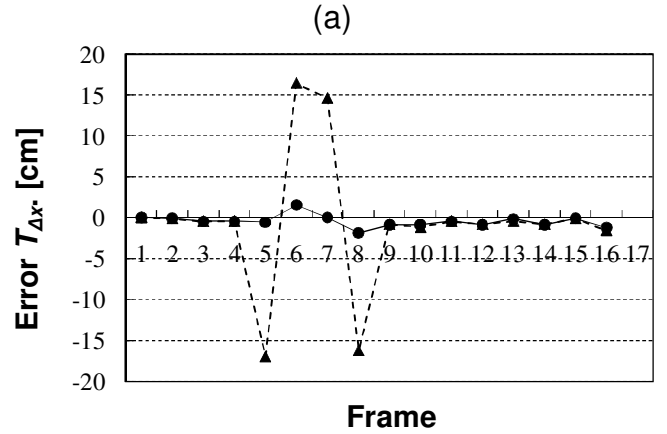**Figure 6.** Experiment 2. (a) top view (b) side view (c) before (d) after.

Figure 7. Errors in the experiment 2. (a) x direction (b) y direction.

Example 2

Figure 6a and 6b illustrate top view and side view of this experiment respectively. The debris brings about the error $T_{\Delta x}$ and also $T_{\Delta y}$. Figure 6c and 6d illustrate the estimated locations on a ceiling map before and after applying the proposed method respectively. In Figure 6d, locations are plotted in a line at a constant interval as the errors are eliminated.

Figure 7a and 7b summarizes the error $T_{\Delta x}$ and $T_{\Delta y}$ for the case in Figure 6. It is observed that the maximum absolute value of the error is reduced from 17 to 2 cm and from 9 to 1 cm respectively. It is observed that the proposed method can reduce the error to 11.76 and 11.11% respectively. It is confirmed that utilization of the forward camera's video reduced the error to approximately 11%.

Conclusions

In this paper, we reduced errors in estimating location of an indoor mobile robot on a ceiling map by utilizing visual

error is reduced from 21 to 2 cm. It is confirmed that the proposed method can reduce the error to 9.52%.

compensation of motions in the "upward" camera's video with those in the "forward" camera's video. It was experimentally confirmed that the maximum absolute value of the error was reduced to approximately 11%. Future work such as modification of the current mathematical model or other approach will be done to enhance current estimation performance.

REFERENCES

Desouza GN, Kak AC (2002). Vision for Mobile Robot Navigation: A Survey. IEEE Trans. Pattern Anal. Machine Intell., 24(2): 237-267.
Ito K, Nakajima H, Kobayashi K, Aoki T, Higuchi T (2004). A Fingerprint Matching Algorithm Using Phase-Only Correlation. IEICE Trans. Fundam., E87-A(3): 682-691.
Munguia R, Grau A (2007). Monocular SLAM for Visual Odometry. IEEE International Symposium on Intelligent Signal Processing (WISP), pp. 1-6.
Nister D, Naroditsky O, Bergen J (2004). Visual Odometry. Proc. IEEE Computer Soc. Conf. Computer Vision Pattern Recognition (CVPR), 1(27): I-652-I-659.
Papanikolopoulos NP, Khosla PK, Kanade T (1993). Visual tracking of a moving target by a camera mounted on a robot: A combination of control and vision. IEEE Trans. Robot. Automation, 9(1): 14-35.
Sasaki H, Kobayashi K, Aoki T, Kawamata M, Higuchi T (1998). Rotation measurements using rotation invariant phase-only correlation. ITE Tech. Rep., 45: 55-60.
Udomsiri S, Taguchi H, Takahashi T, Iwahashi M, Kimura T (2009). Functionally Layered Video Coding Based on JP2K for Robot Vision Network. J. Robot. Mechatronics, 21: 6.
Wilson CA, Theriot JA (2006). A Correlation-Based Approach to Calculate Rotation and Translation of Moving Cells. IEEE Trans. Image Process., 15(7): 1939-1951.