BAND DETERMINATION OF JP2K DWT FOR ROBOT TO ROBOT COMMUNICATION

Tomoyuki TAKAHASHI   Yasuaki FUKUTANI   Masahiro IWAHASHI

Depertment of Electrical Eng.,
Nagaoka University of Technology,
Nagaoka-shi Niigata, 940-2188, JAPAN
E-mail: ttakahashi@tech.nagaokaut.ac.jp

ABSTRACT

The system in this paper generates a ceiling map from video obtained by an upward camera of a moving indoor robot. The map is shared by robots for auto-localization. When the floor is not flat, automatically estimated location contains error. We compensate the error by utilizing video from a forward camera of the robot. In the system, a remote server estimates the location based on data transmitted by a robot. Therefore it requires doubled data amount for communication to add a forward camera. We reduce total data to be transmitted by removing redundancy of motion information between videos from an upward camera and a forward camera.

Keywords: robot, vision, network, coding, JPEG2000

1. INTRODUCTION

Auto-localization of a moving robot is a challenging problem for a robot vision network, in which various kinds of moving (or mobile) robots communicate and cooperate automatically. In case of a global positioning system (GPS) can't be used, acceleration sensor, geomagnetic attitude sensor, range sensor and gyroscope are useful tools for this purpose.

Comparing to these sensors, video based localization methods are less accurate and robust in general. However, it is still attractive for constructing a small sized energy saving moving robot with micro CCD sensors. It is becoming feasible owing to recently developed odometry and simultaneous localization and mapping (SLAM) techniques [1-3].

In this paper, we deal with a system in which a ceiling map is generated from video signal obtained by an upward camera of a moving indoor robot. The map is shared by other moving robots plotting their locations on the same map [4]. In the system, a remote server estimates their locations based on data transmitted by each robot.

There are two important problems to be solved in such system. One is to estimate precise location in the map from video. The other is to compress data amount to be transmitted by the robots via digital network.

Correlation based matching between adjacent two frames of video is one of successfully utilized techniques for map generation and location estimation [5]. However it is sensitive to lighting conditions. Recently, phase only correlation (POC) and rotation invariant POC (RI-POC) have been proposed for stable motion estimation [6,7].

In our previous work, we confirmed that POC is robust to spot light disturbance and useful for the system [8]. We also reduced data amount to be transmitted to the remote server from a moving robot by removing redundancy in video for motion estimation [9].

However, when a robot moves on the floor not flat with debris, automatically estimated location from an upward camera of the robot becomes imprecise. This is the problem to be solved in this paper.

Firstly we improve accuracy of location estimation by compensating motion information from an upward camera by that from a forward camera. If the floor is flat, the compensation is not necessary. However, on the non-flat floor, it reduces error in motion information from the upward camera.

Since we add data from one more camera, it requires doubled data amount to be transmitted in general. Therefore, secondly we reduce total data for communication between a moving robot and a remote server by removing redundancy of motion information between videos from an upward camera and a forward camera.

A moving robot of the proposed system has a couple of upward and forward CCD image sensors. It also includes compression algorithm same as the JPEG 2000 (JP2K) international standard [10,11]. It just picks up the minimum components of wavelet transform coefficients and bit-planes for transmission to a remote server for auto-localization. It is also possible to see video scenery. It is expected to contribute for precise location estimation and also low bit rate data transmission.

2. EXISTING METHOD

The auto-localization system based on a ceiling map to be discussed in this paper is summarized. Its problem on accuracy of location estimation on non-flat floor is addressed. Existing data compression method using upward camera only is summarized.

2.1 Auto-Localization in Ceiling Map

Fig.1 illustrates an example of a ceiling map. It is generated from video signal of an upward camera mounted on a moving robot. Thirty frames per second are connected each other after adjusting their positions. This adjustment is based on motion information estimated with RI-POC.

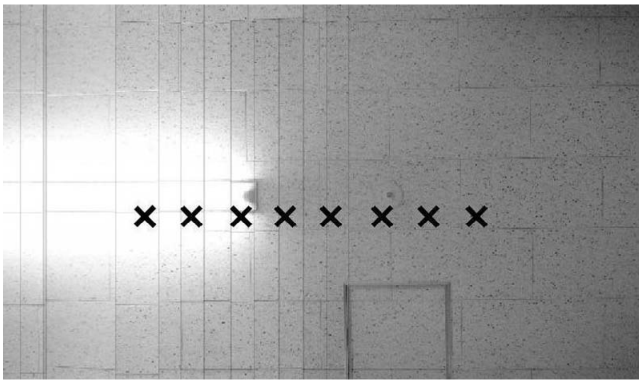

Fig.1 Ceiling map generated from an upward camera of a robot moving on a flat floor.

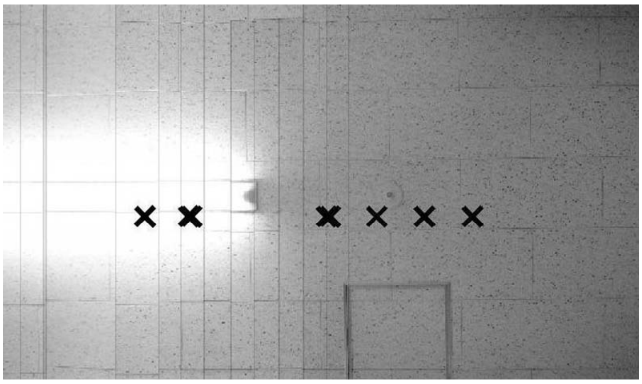

Fig.2 Ceiling map generated from an upward camera of a robot moving on a non-flat floor.

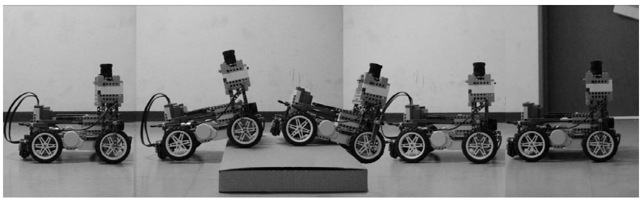

Fig.3 A robot moving on a non-flat floor.

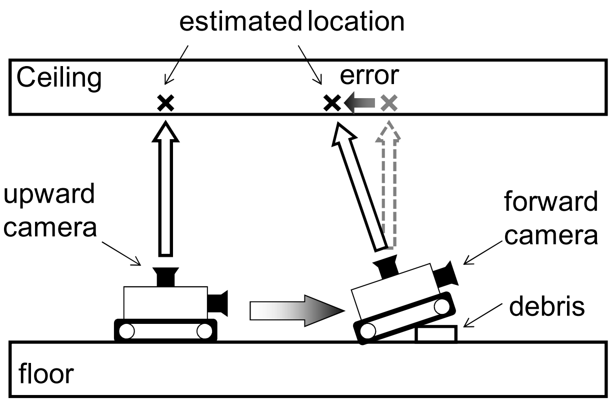

Fig.4 Location estimated from an upward camera of a robot contains error due to debris on the floor.

In the figure, a "cross" denotes location of a robot at an equal time-interval. The crosses are plotted at an interval on the ceiling map since the robot moves on a flat floor at the same speed. Once the map is generated, locations of other robots can be similarly plotted. It is expected to be utilized for cooperative recur action by plural moving indoor robots.

Fig.2 illustrates a map generated from an upward camera of a robot moving on a non-flat floor. This situation is illustrated in Fig.3. Since the robot moves at a constant speed, the cross should be plotted at an interval as in Fig.1. However, in this case, estimated location contains error due to debris on the floor.

As illustrated in Fig.4, it is inevitable to have error in motion information from an upward camera only. We compensate this error by introducing motion information from a forward camera as described in **3.1**.

2.2 Minimum Data Transmission

Fig.5(a) illustrates the JP2K international standard for video data compression. It transmits the same data to a user and a remote server. Data amount of video is compressed, however the remote server uses video only for estimation of motion information.

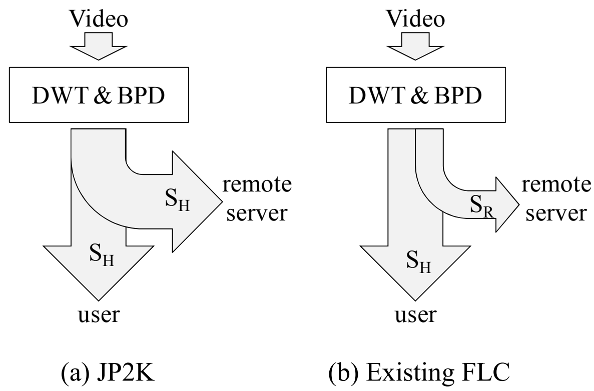

Fig.5 The existing FLC in [8] transmits the minimum data to a remote server necessary for motion estimation.

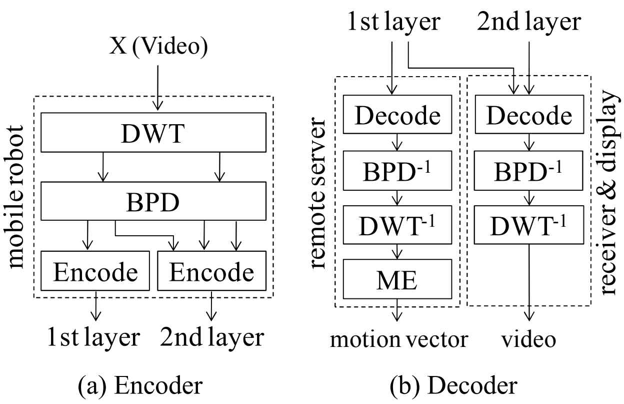

Fig.6 Encoder and decoder of the existing FLC in [8]. ME: motion estimation, BPD: bit plane decomposition, DWT: discrete wavelet transform.

Fig.5(b) illustrates the existing functionally layered coding (FLC) [8]. It transmits the minimum data to a remote server necessary for motion estimation (ME). Therefore data amount for communication between a robot and a remote server is reduced.

Fig.6(a) illustrates encoder of FLC. Discrete wavelet transform (DWT) and bit plane decomposition (BPD) decompose video into several band and bit plane components. It is encoded with EBCOT in the same manner of JP2K. The minimum components for ME are embedded in the 1st layer for a remote server.

Fig.6(b) illustrates decoder. RI-POC in ME estimates three parameters: 1) rotation angle, 2) scaling (dilation) rate and 3) translation (motion vector) as motion information. POC estimates translation only. Therefore RI-POC needs more components than POC.

The proposed method has two cameras and therefore total data amount for ME are doubled in general. However we remove redundancy of motion information between these two videos to reduce total data as described in **3.2**.

3. PROPOSED METHOD

We add a forward camera to solve accuracy problem of location estimation described in **2.1**. We remove redundancy on motion information between two videos from two cameras to reduce total data described in **2.2**.

3.1 Improvement of Location Estimation

Fig.7 illustrates an analysis model for compensation of motion information obtained from an upward camera by that of a forward camera. Relation among parameters in the figure is described as follows:

$$\begin{aligned} & d-\left(d+a\tan\theta_{fb}\right)\cos\theta_{fb} \\ & +\left\{\left(h+\frac{C_v}{\tan\theta_{fb}}\right)\tan\theta_{fb}-\left(d+a\tan\theta_{fb}\right)\sin\theta_{fb}\right\} \\ & \times\left\{\tan\theta+\frac{\tan\theta_2}{\tan\theta_1-\tan\theta_2}\left(\tan\theta_{fb}-\tan\theta_1\right)\right\} \\ & =0 \end{aligned} \quad (1)$$

According to Eq.(1), up and down translation from the forward video determines the angle θ_{fb} due to pitch of the robot caused by debris. Similarly the angle θ_{rl} due to roll of the robot is determined by rotation angle from the forward video. In this paper we are detecting surge and yaw precisely under roll and pitch neglecting heave and sway.

As a result, the translation from the upward video is compensated by

$$\begin{pmatrix}\Delta x\\ \Delta y\end{pmatrix}=\begin{pmatrix}h\cdot\tan\theta_{rl}-a\cdot\sin\theta_{rl}\tan\theta_{rl}\\ h\cdot\tan\theta_{fb}-a\cdot\sin\theta_{fb}\tan\theta_{fb}\end{pmatrix} \quad (2)$$

where a and h are already known distance illustrated in Fig.7. Eq.(2) solves the auto location problem due to debris on the floor. It compensates motion information from the upward video by that from the forward video.

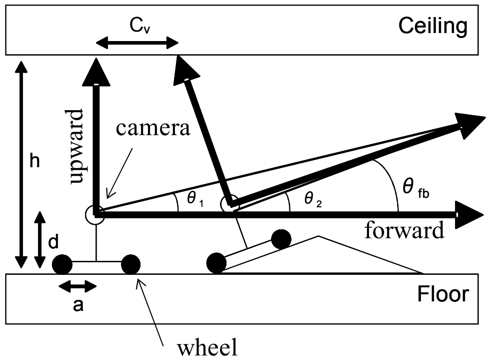

Fig.7 Analysis model for compensation of motion information of upward video by that of forward video.

3.2 Reduction of Transmission Data

In general, data amount to be transmitted is doubled since we add one more video from the forward camera. In this paper, we utilize the fact that rotation angle from the upward camera is not necessary because translation (right and left) from the forward video can be used instead of it. It means that we detect yaw from forward camera (not from upward camera). Therefore we use POC instead of RI-POC for the upward video to extract translation only.

When RI-POC is applied to both of two videos, total data amount of components for ME is doubled. Comparing to this case, our method can reduce total data amount.

4. EXPERIMENTS

Effectiveness of the proposed method is experimentally confirmed evaluating performance of auto localization and data compression.

4.1 Error in Auto-Localization

Fig.8 summarizes experimentally evaluated error in location estimated for the case in Fig.3. The angle θ_{fb} is measured in the range of [-4, +4] [degree]. As indicated in the figure, the error was not compensated in the existing method without utilizing the forward video. The error reached 20.5 [cm]. In this case, a generated ceiling map is indicated in Fig.2.

In the proposed method, error was found to be 1.4 [cm] at most. A generated map was close to (almost the same as) Fig.1. Effectiveness of the utilization of motion information from the forward camera was confirmed.

4.2 Data Amount to be Transmitted

Table 1 and table 2 summarize error of rotation angle and translation respectively. These are estimated with RI-POC. Table 3 summarizes error of translation estimated with POC. In the experiment, two image areas with 256x256 pixels at different random positions in a ceiling image were used. The tables indicate average over 100 trials. The maximum of translation and rotation were set to 30 [pixel] and 10 [degree] respectively.

For example, in table 1, (all, 9) means the case where all the bands and 9 bit planes are used for ME. In this case, error average is 0.16 [degree] and 243.61 [kB] data amount is necessary to be transmitted according to table 4. When only LL band in 1 stage octave decomposition is used [9], error increased to 0.29 [degree] but data size decreased to 66.74 [kB]. Namely, accuracy and data size are in trade off.

For example, when a tolerable error is set to 1 [pixel] in the ceiling map, it is equivalent to 0.44 [degree] in average, RI-POC requires (1LL, 8) with 46.73 [kB] to transmit at minimum bit rate. On the contrary, POC requires (1LL, 6) with 10.82 [kB].

It means that the existing method needs 46.73 [kB] for RI-POC on each of the upward and forward video. However, the proposed method needs 46.73 [kB] for RI-POC on the forward and 10.82 [kB] for POC on the upward video.

The existing method requires 93.46 [kB] whereas the proposed requires only 57.55 [kB] in total for transmission to a remote server. It was confirmed that the proposed method reduces data amount to be transmitted to 61.6 [%].

5. CONCLUSIONS

We compensated location estimation error in ceiling map by utilizing video from a forward camera of a robot and confirmed its improvement. We also reduced total transmission data to 61.6 [%] by removing redundancy of motion information between videos from an upward camera and a forward camera.

Table 1 Error of rotation angle of RI-POC in [degree].

bit plane		9	8	7	6	5
band	all	0.16	0.15	0.63	6.46	12.56
	1LL	0.29	0.30	3.00	8.87	12.71
	2LL	5.03	5.47	8.59	9.50	10.61
	3LL	6.44	5.99	8.02	9.27	-

Table 2 Error of translation of RI-POC in [pixel].

bit plane		9	8	7	6	5
band	all	0.00	0.00	0.25	13.07	29.83
	1LL	0.24	0.23	7.04	19.93	30.79
	2LL	15.77	15.36	20.96	25.13	35.73
	3LL	25.68	29.04	34.58	32.43	-

Table 3 Error of translation of POC in [pixel].

bit plane		9	8	7	6	5
band	all	0.00	0.00	0.00	0.00	3.81
	1LL	0.73	0.73	0.73	0.73	15.17
	2LL	1.16	1.15	1.14	4.69	28.92
	3LL	4.96	8.24	14.72	26.02	-

Table 4 Data amount to be transmitted in [kB].

bit plane		9	8	7	6	5
band	all	243.61	163.78	86.88	31.34	6.10
	1LL	66.74	46.73	26.81	10.82	2.71
	2LL	17.61	12.40	7.12	2.96	1.09
	3LL	4.97	3.59	2.22	1.24	0.83

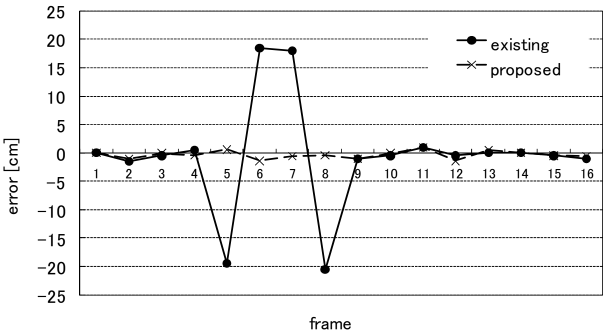

Fig.8 Evaluation results of error in location estimation.

6. REFERENCES

[1] G. N. Desouza, A. C. Kak, "Vision for Mobile Robot Navigation: a Survey", IEEE Trans. Pattern Analysis Machine Intelligence, vol.24, Issue 2, pp.237-267, Feb. 2002.

[2] D. Nister, O. Naroditsky, J. Bergen, " Visual Odometry ", Proc. IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), Vol. 1, 27, pp. I-652 - I-659, July 2004.

[3] R. Munguia, A. Grau, "Monocular SLAM for Visual Odometry", IEEE International Symposium Intelligent Signal Processing (WISP), pp.1-6, Oct. 2007.

[4] N. P. Papanikolopoulos, P. K. Khosla, T. Kanade, "Visual tracking of a moving target by a camera mounted on a robot: a combination of control and vision", IEEE Trans. Robotics Automation, Vol.9, Issue1, pp.14-35, Feb. 1993.

[5] C. A. Wilson, J. A. Theriot, "A Correlation-Based Approach to Calculate Rotation and Translation of Moving Cells", IEEE Trans. Image Processing, Vol.15, Issue 7, pp.1939-1951, July 2006.

[6] K. Ito, H. Nakajima, K. Kobayashi, T. Aoki, T. Higuchi, "A Fingerprint Matching Algorithm Using Phase-Only Correlation", IEICE Trans. Fundamentals, Vol.E87-A No.3 pp.682-691, March 2004.

[7] H. Sasaki, K. Kobayashi, T. Aoki, M. Kawamata, T. Higuchi "Rotation measurements using rotation invariant phase-only correlation" ITE Tech Rep 45: 55-60, 1998.

[8] M. Iwahashi, H. Taguchi, T. Kimura, "Functionally Layered Coding for Robot Vision Network", Picture Coding Symposium, no.2, pp.32-35, May 2009.

[9] S. Udomsiri, H. Taguchi, T. Takahashi, M. Iwahashi, T. Kimura, "Functionally Layered Video Coding Based on JP2K for Robot Vision Network", Journal of Robotics and Mechatronics Vol.21 No.6, Dec. 2009.

[10] JTC1/ SC29, "Information technology - JPEG 2000 image coding system: Core coding system", ISO/IEC15444-1, 2004.

[11] A. Descampe, F. O. Devaux, G. Rouvroy, J. D. Legat, J. J. Quisquater, B. Macq, "A Flexible Hardware JPEG 2000 Decoder for Digital Cinema", IEEE Trans. CAS for Video Technology, vol. 16, issue 11, pp.1397 - 1410, Nov. 2006.