VISUAL ODOMETRY OF A MOBILE PALETTE ROBOT USING GROUND PLANE IMAGE FROM A FISHEYE CAMERA

Ung-Gyo Lee[1], Soon-Yong Park [2] *

[1] School of Robot and Smart System Engineering, Kyungpook National University, Deagu, South Korea, dndry123@naver.com
[2] School of Electronic and Electrical Engineering, Kyungpook National University, Deagu, South Korea, sypark@knu.ac.kr

Commission I, ICWG I/IV

KEY WORDS: Visual Odometry, Palette Robot, Orthophoto, Fisheye Camera, Kinematic Model, Ground Plane Image

ABSTRACT:

In this paper, we present a method of mobile robot's visual odometry using the visual feature tracking in the ground plane image generated from the fisheye image. In order to extract the feature information on the ground, we use a fisheye camera that has a larger FOV than a general pinhole camera, so that we can capture more information on the ground plane. However, due to the large distortion, it is difficult to extract the visual features in the fisheye image. The distortion can be eliminated, but various problems arise, such as a decrease in the resolution of an image or losing the wide angle of the fisheye camera. We propose the EUCM-Cubemap projection model to convert the fisheye image into the cubemap image without losing the FOV of the fisheye image. And we create the Ground Plane Image, a virtual image that vertically looks at the ground from a cube map image. So, the ground plane image is generated so that it is captured from a virtual camera perpendicular to the ground. In the ground plane image, the motion vector obtained by feature tracking between previous and current frames is proportional to the actual robot's motion in the 2D ground plane. Thus, if we know the actual scale of the motion vector, we can estimate the mobile robot's velocity and steering angle on the virtual wheel generated by the ground plane image. The scale of the vector can be estimated using the position and focal length of the camera. Using these parameters, we estimate the mobile robot's pose by applying the bicycle kinematic model. Experimental results show that the proposed method can replace other conventional odometry methods for mobile robots. And, in the future, it is expected to be used in a variety of fields such as visual-based control or path planning.

1. INTRODUCTION

Today with the increased integration of chipsets and the development of efficient structures, the computing technology of mobile computers has made rapid progress compared with the past. This has led to the development of autonomous vehicles, service robots, and industrial robots. Because, in the past, technologies could not be applied due to limitations in hardware performance. And now, we can apply to such fields that require complex epistemological algorithms in miniaturized computers such as Robotics, Augmented Reality, and Virtual Reality. Computer vision technology has expanded its field further as the evolutionary leap of AI technology and so, it is being used in various fields, including manufacturing, national defense industry, and even service industries. Especially, in such fields as autonomous vehicles or mobile robots, it is important to minimize blind spots by expanding the perception range, so a lot of sensors should be installed. The camera sensors are usually used because cameras are cheap, readily available, and can be used for general purposes in a wide field. Moreover, by using a fisheye camera that has a large FOV, the perception range can be expanded with the minimum number of cameras. Because, in a fisheye camera, the amount of large information is entered in a single frame and it is effective to reduce costs. So, the importance of a fisheye camera is rapidly growing up. The localization of a mobile robot is an essential technology in autonomous driving. However, such as palette robots and cleaning robots, a camera is attached close to the

ground. In these environments, camera-based localization is a challenging problem because feature tracking on random patterns in variable light environments with high-frequency vibrations results in less accuracy. Another solution is to use Wheel Odometry which estimates the mobile robot's location by attaching an encoder to the robot's wheel, but there are limitations such as the wheel's slip or backless and caused by different friction depending on what material the floor is. Also, the vibration causes the divergence of IMU which significantly lowers the accuracy depending on the specific environment. In this paper, we propose a method of applying the information on the ground obtained from the fisheye camera, we estimate the velocity and steering angle of the virtual wheel in the image and apply the kinematic model of the mobile robot to localization. The proposed method is largely composed of three parts. First, a fisheye image is conceived as a cubemap image, and a virtual image Ground Plane Image perpendicular to the floor surface is generated from the conceived cube image. And, as shown in Fig. 1, in this image, feature points are Optical tracked to obtain a vector generated by movement in the image, and the location of the robot is estimated using the kinematic model of the mobile robot. The kinematic model used in this paper used the bicycle motion model. The structure of our paper is as follows. We provide five chapters: Chapter 2, briefly introduces similar studies related to this paper. And in Chapter 3, we explain how to make up each system of proposed methods. And Chapter 4 shows the experiments. Finally, in Chapter 5, we describe the conclusions.
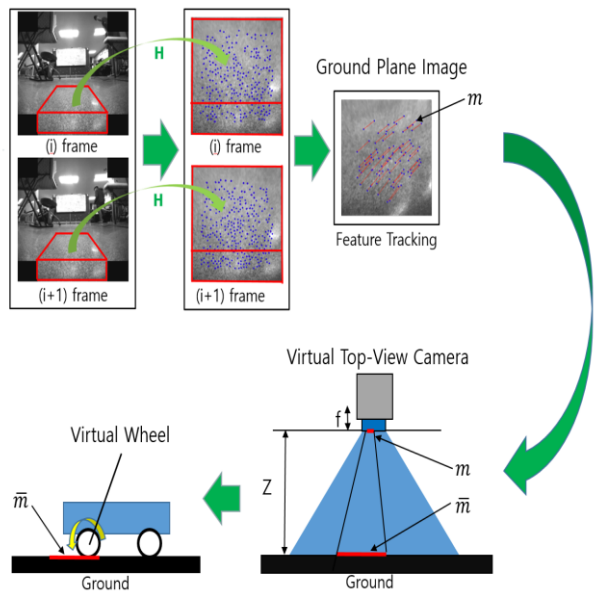
---

* Corresponding author

Figure 1. Belief illustration of our visual odometry system.

2. RELATIVE WORK

In this section, we briefly introduce the previous studies which are related to mobile vehicle a robot's visual odometry, using the features on the surface of the ground and a fisheye image.

2.1 Cube SLAM: A Piecewise-Pinhole Monocular Fisheye SLAM System

This work presents a novel CubemapSLAM system that incorporates the cubemap model into the ORB-SLAM which is a state-of-the-art feature-based SLAM system. The cubemap model utilizes a large FoV of the fisheye camera without affecting the performance of feature descriptors. In addition, CubemapSLAM is efficiently implemented and can run in real-time. In this paper, despite the limited angular resolution of the sensor, CubemapSLAM shows better accuray than pinhole camera projection model. (Wang, Yahui et al., 2018)

2.2 Localization using visual odometry and a single downward-pointing camera

This work demonstrates the use of a single downward-pointing camera and visual odometry techniques for localization. The technique uses feature detection and optical flow measurements to provide sensor information to localization algorithms. In this paper, the application is specifically targeted to robotic platforms in unknown areas such as GPS-denied and barren environments. (Swank, Aaron J. et al., 2012)

2.3 Robust monocular visual odometry for a ground vehicle in undulating terrain

This work presents a robust method for monocular visual odometry capable of accurate position estimation even when operating in undulating terrain. In this paper, using a steering model to separately recover rotation and translation. And proposed method handles undulating terrain by approximating ground patches as locally flat but not necessarily level, and recovers the inclination angle of the local ground in motion estimation. In the field, experiments show an error of less than 1%. (Zhang, Ji et al., 2014)

2.4 Kinematic model based visual odometry for differential drive vehicles

This paper presents the visual odometry of a vehicle that is working on the two-dimensional plane by applying a mobile robot's kinematic model using a monocular camera that is facing the floor. The system is inexpensive and efficient enough to be used in real time on a single CPU. (Jordan, Julian et al., 2017)

3. APPROACHS

3.1 System Overview

A flow diagram of the proposed method is shown in Fig. 2. First, we convert the fisheye image into the cubemap image. From the cubemap image, we create a virtual camera image called Ground Plane Image (GPI) that captures the ground from the vertical direction from the camera. The GPI is generated from the cubemap's Bottom-Face and Front-Face images as shown in Fig. 6. And as shown in Fig. 7, to find the robot's motion, we track feature points in GPI using an optical flow algorithm. We use the KLT (Kanade-Lucas-Tomashi) algorithm to find the motion between consecutive two frames. After calculating the actual scale of the vector, the mobile robot's velocity and steering angle are estimated. For accurate motion estimation, we exploit the robot's kinematic model for odometry. In this paper, we use the bicycle motion model (Polack, Philip et al., 2017) for our mobile palette robot. This model represents a wheel-based mobile robot's motion. As shown in Fig. 9, we estimate the mobile robot's odometry by applying the kinematic model.
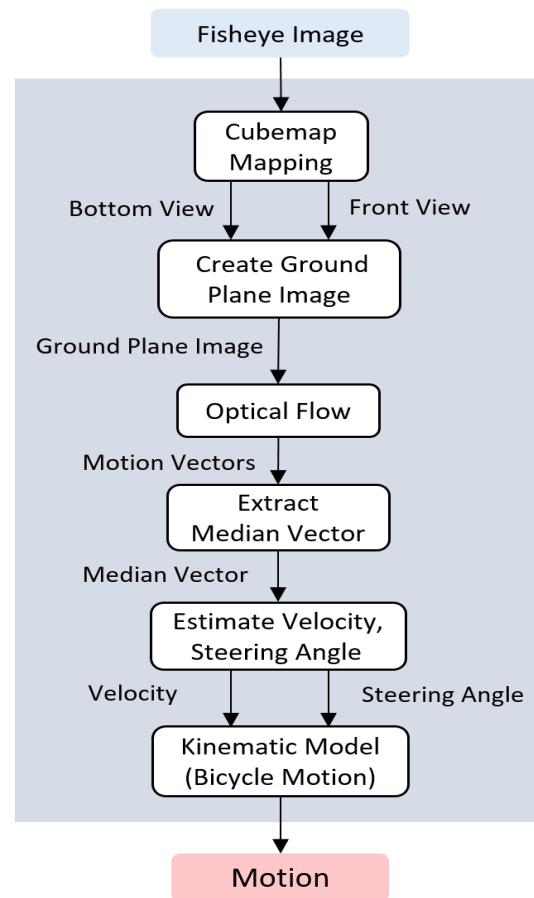
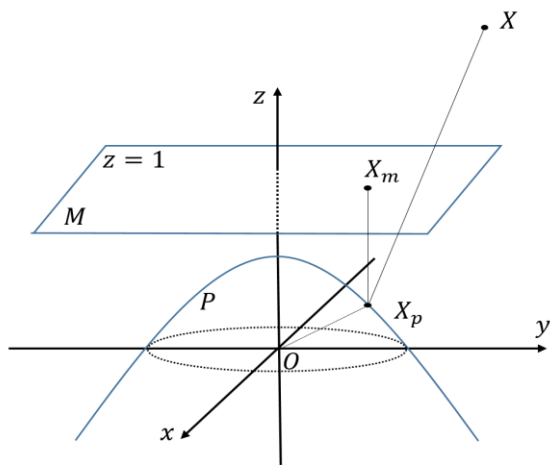

Figure 2. Proposed total system diagram.

Figure 3. Illustration of Enhanced Unified Camera Model.

3.2 Calibration Fisheye Camera Using EUCM

When using the fisheye image, applying the pinhole projection camera model causes serious distortion in the undistorted image. In addition, it causes serious raster distortions in the undistorted image. Therefore, instead of pin-hole model, we use a fisheye calibration method, EUCM (Enhanced Unified Camera Model) (Khomutenko et al., 2015) which is a nonlinear camera model that is generally used in omnidirectional camera such as catadioptric systems or fisheye cameras. Explaining the EUCM briefly, EUCM has more parameters than the pinhole model's parameters. So, using these parameters the EUCM can represent the camera distortion without any undistort process. As shown in Fig. 3 point X is projected on the curve plane P and then projecting it orthogonally into the M plane which is the normal plane. In this paper, to calibrate the fisheye camera, we used a calibration toolbox called Kalibr. It is a very useful tool for camera calibration. The Kalibr offers a variety of camera models including EUCM, Pinhole Model, Omnidirectional Model and Double Sphere Model.

3.3 Mapping Fisheye Image to Cubemap Image

After fisheye camera calibration, we convert the fisheye image to the cubemap image. As shown in Fig. 4, the cubemap model consists of five image planes generated by virtual pinhole projection model. Each virtual pinhole model has the same intrinsic parameters but has different extrinsic parameters to look in different directions, so it is possible to divide a large FOV image into multiple virtual pinhole cameras. As shown in Fig. 5, (b) fisheye image is mapped in a cube model that includes a virtual pinhole model on each face.
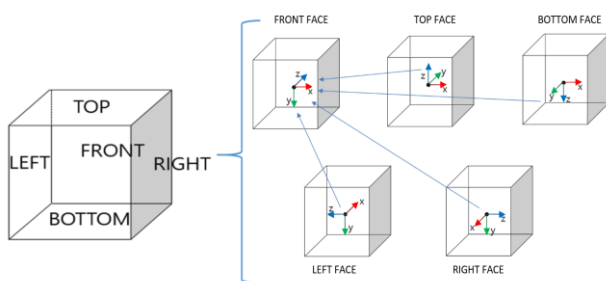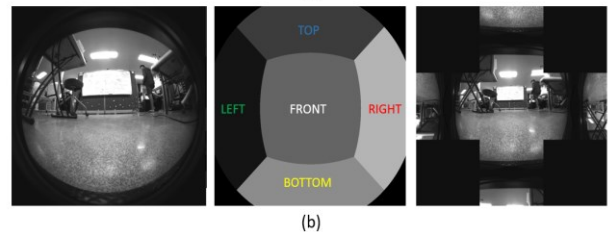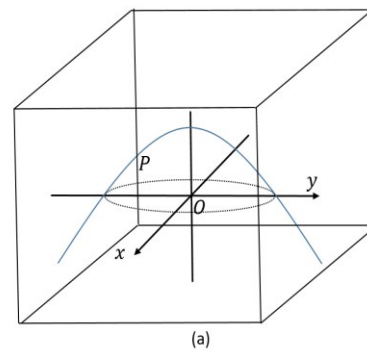

Figure 4. Illustration of Cubemap Model.

Figure 5. (a) Illustration of EUCM-Cubemap projection model. (b) Convert fisheye image to the cubemap Image.

3.4 Generating Ground Plane Image from Cubemap Image

As shown in Fig. 6, the Front and Bottom Faces in the cubemap image are used to generate the Ground Plane Image. In the Front-Face, a homography matrix warps the Front-Face image to the Ground Plane Image. And the Bottom-Face image in the cubemap image is used as it is because the Bottom-Face is already obtained by a virtual top view camera in the cubemap model as shown in Fig. 5 (a). We assume that this image is a virtual orthophoto, which looks vertically at the ground and called Ground Plane Image.

3.5 Feature Tracking on Ground Plane Image

In order to obtain a motion vector generated by the mobile robot's movement, the vector is represented in the Ground Plane Image because it is a virtual image perpendicular to the floor surface generated by cubemap image. And then, perform the feature tracking between the previous image and the current image. In this paper, we used KLT (Kanade-Lucas-Tomasi) optical flow algorithm for the feature tracking on the Ground Plane Image. And, in order to obtain a reliable optical vector that represents the mobile robot's motion in the Ground Plane Image. As shown in Fig. 7, we calculate the median vector of motion vectors on the Ground Plane Image and assume that the median vector is reliable and calculate a median vector among the motion vectors.
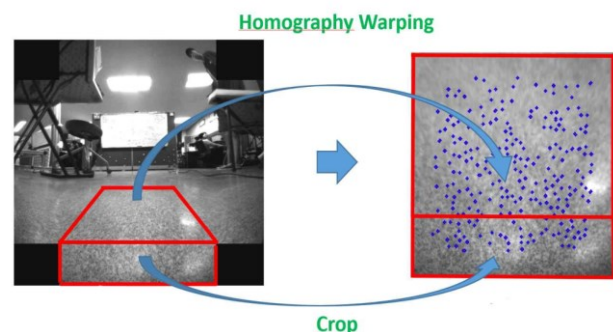

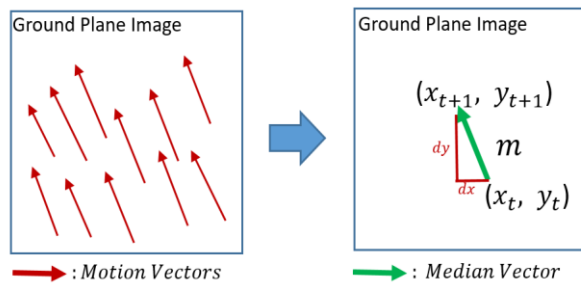**Figure 6**. Generate Ground Plane Image from the cube image.

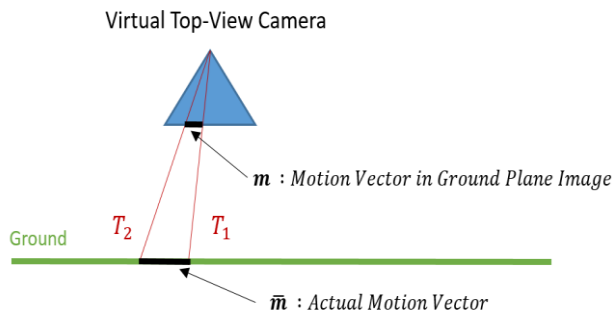**Figure 7**. The motion vectors on the Ground Plane Image.

Figure 8. Illustration of motion vector in Ground Plane Image and actual motion vector

3.6 Motion Vector on Ground Plane Odometry

As shown in Fig. 8, the motion vectors obtained from the Ground Plane Image are matched with the mobile robot's motion with an unknown scale. Because the location of the camera is depending on the robot's motion. Thus, if we know the real scale of the motion vector, we can estimate the actual motion of the robot in the Ground Plane Image. To estimate the scale of the motion vector, some assumptions are followed. Assuming that the Ground Plane Image is obtained by a virtual camera which is perfectly perpendicular to the flat ground, the scale of the motion vector is estimated using the height from the ground to the camera mounted on the mobile robot and the focal length of the camera. In this case, the motion vector obtained from the Ground Plane Image multiplied by the scale represents the mobile robot's motion and the camera in two dimensions.

3.7 Visual Odometry using the Bicycle Motion Model

The velocity and the steering angle of the mobile robot's virtual wheel are calculated by using the scaled motion vector obtained from Ground Plane Image. Using these measurements, we apply the virtual wheel's measurements to the mobile robot's motion model. As shown in Fig. 9, the estimated values of Speed V and String Angle δ of the virtual wheel are measured by the scaling motion vector obtained from the Ground Plane Image, and wheel odometry can be performed by using visual information in a two-dimensional space. As shown in Fig. 9, the bicycle motion model is a Wheel-Based vehicle's kinematic model that rotates and moves based on the center point IC (Instantaneous Centre) in a two-dimensional plane and estimates the mobile robot's odometry. The L is the distance between the mobile robot's rear-wheel axis and front-wheel axis, the lr is the distance between the rear-wheel axis and the center of mass of the mobile robot. The w is the angular velocity at which the mobile robot moves based on the center of the IC.
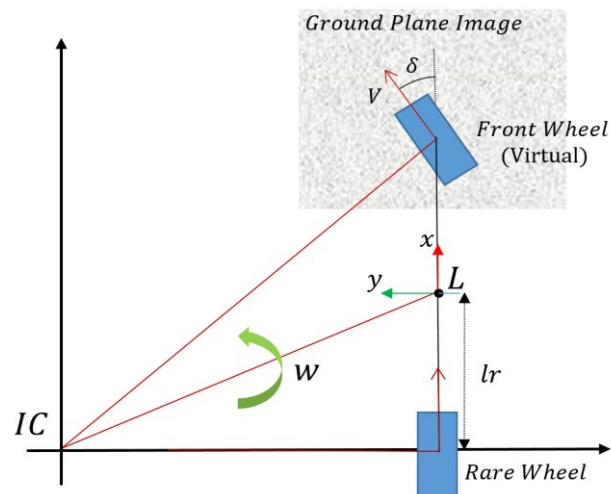

Figure 9. Motion vector on the Ground Plane Image applying to the mobile robot's motion model

4. EXPERIMENTS

4.1 Configuration and environment

The robot used in the experiment was performed by our ROS-based pallet robot (Lee, Ung-Gyo et al., 2021) Table.1 shows the mobile robot's hardware spec and as shown in Fig. 10 (b) the fisheye camera is mounted on the head of the Mobile Pallet Robot. In the experiment, the camera we used is Intel's Tracking Camera T265. T265 is a fisheye stereo camera. It has two fisheye lenses with a combined amount of 163±5 Field of View, and it has (848x800) resolution for each camera lens and also, has an IMU because T265 performs localization based on VIO (Visual Initial Odometry).

Component	Hardware
Embedded Board	NVIDIA Jetson-Nano
Wheel Encoder	PHIDGET Optical Rotary Encoder
Motor	PHIDGET NEMA-17 Bipolar Stepper
Motor Controller	PHIDGET Stepper Bipolar HC
LiDAR	SLAMTEC RPLIDAR-S1
Camera	Intel RealSense T265 Tracking Camera
Battery	NEXT-408PB-UPS(40800maAh)

Table 1. Pallet robot's hardware spec sheet
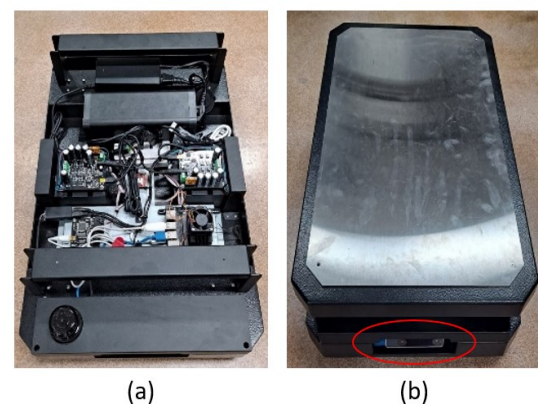

(a)   (b)

Figure 10. (a) Our pallet robot (b) A camera attached in front of the robot

4.2 Experiment

A camera is placed on a mobile robot that is attached to the floor as shown below, and when looking at the image, an image in which the floor occupies more than half of the image is obtained. The real-world experiments are performed using a mobile palette robot which is equipped with Intel Tracking Camera T265 as shown in Fig. 10 (b). In the experiment, Fig. 11 (a) shows unfiltered motion vectors that a representative measurements of optical flow on Ground Plane Image. Fig. 11 (b) represented the distribution of motion vectors and (c) shows a histogram generated by the motion vector's angles in the Ground Plane Image.
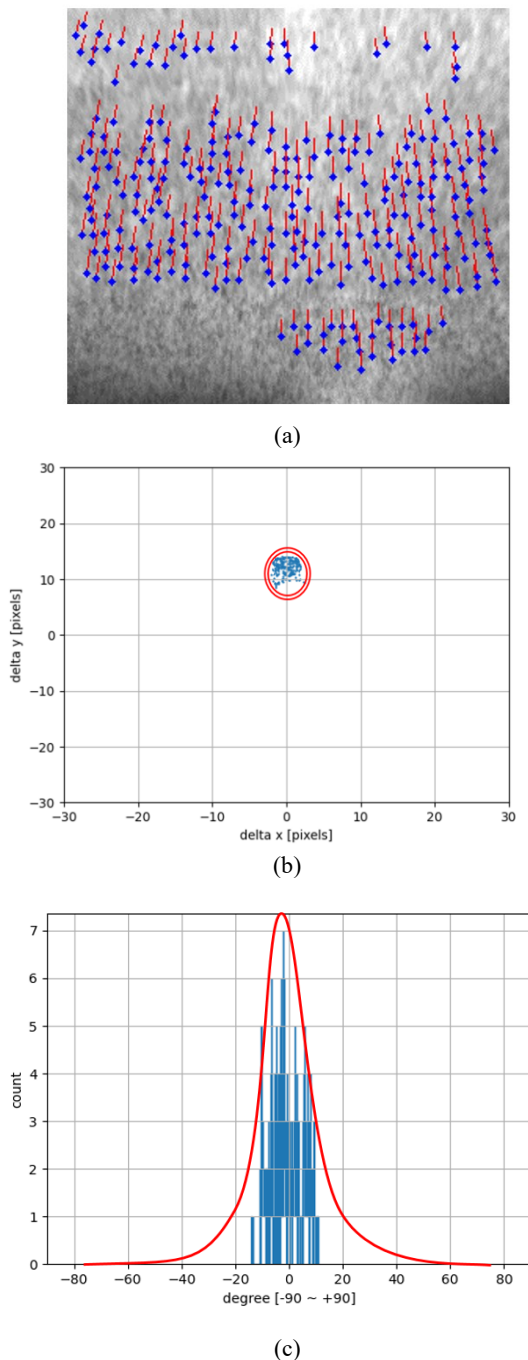

(a)

(b)

(c)

Figure 11. (a) Unfiltered motion vectors obtained from the Ground Plane Image (b) An x-y plot of motion vectors. (c) The histogram on angles of motion vector
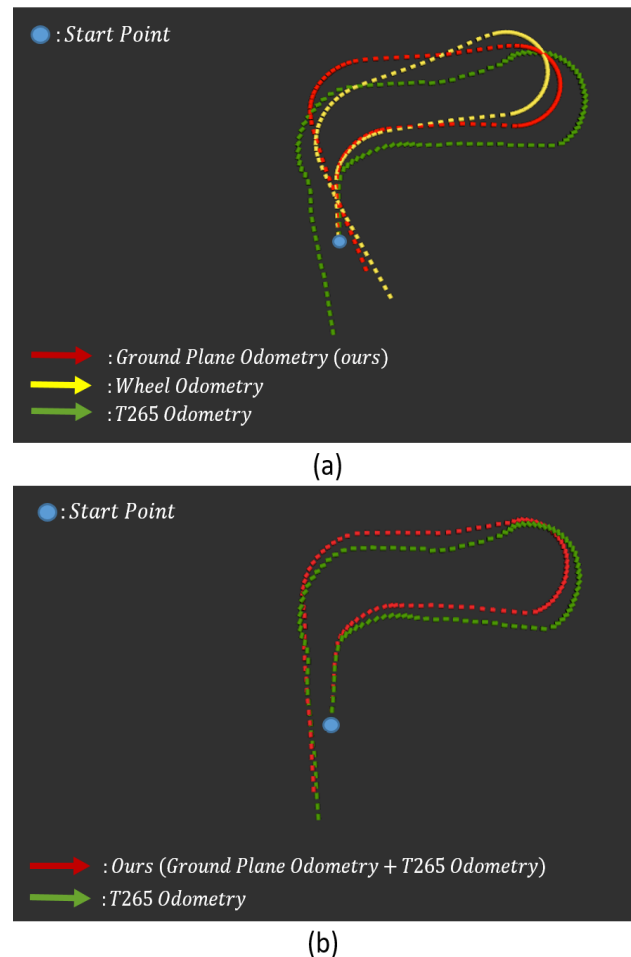

(a)

(b)

Figure 12. (a) Compare odometry (b) Trajectory of proposed method

Fig. 12 (a) shows a comparison of wheel odometry using the mobile robot's wheel encoder, T265 odometry, and the proposed visual odometry called Ground Plane Odometry using the T265 fisheye image. Fig. 12 (b) represents the trajectory of the proposed method by using the EKF(Extended Kalman Filter) to fuse Ground Plane Odometry and T265 Odometry, it is expected that robust VIO localization even in an environment with sparse features and high frequency vibrations.

5. CONCLUSIONS

In this paper, we propose an efficient method of mobile robot's visual odometry with a fisheye camera which is a large FOV camera. In order to use the fisheye image, we mapped a fisheye image to the cubemap image using EUCM and Cubemap Model. Using the cubemap image generate a virtual orthophoto image called Ground Plane Image. Then, the motion vector representing the movement of the robot is obtained from the Ground Plane Image and assumes that it is the measurement of a virtual wheel to apply the mobile robot's motion model and estimate the mobile robot's odometry. In the future, it will be expected that even robots without wheel encoders will be able to wheel odometry using only image information. Furthermore, it is also expected to be used in the field of visual-based control or path planning in the future.

ACKNOWLEDGEMENTS

REFERENCES

Wang, Yahui, et al. "Cubemapslam: A piecewise-pinhole monocular fisheye slam system." *Asian Conference on Computer Vision*. Springer, Cham, 2018.

Khomutenko, Bogdan, Gaëtan Garcia, and Philippe Martinet. "An enhanced unified camera model." *IEEE Robotics and Automation Letters* 1.1 (2015): 137-144.

Polack, Philip, et al. "The kinematic bicycle model: A consistent model for planning feasible trajectories for autonomous vehicles?." *2017 IEEE intelligent vehicles symposium (IV)*. IEEE, 2017.

Swank, Aaron J. *Localization using visual odometry and a single downward-pointing camera*. No. NASA/TM-2012-216043. 2012.

Jordan, Julian, and Andreas Zell. "Kinematic model based visual odometry for differential drive vehicles." *2017 European Conference on Mobile Robots (ECMR)*. IEEE, 2017.

Lee, Ung-Gyo, Kyung-Jea Choi, and Soon-Yong Park. "The Design and Implementation of Autonomous Driving Pallet Robot System using ROS." *2021 Twelfth International Conference on Ubiquitous and Future Networks (ICUFN)*. IEEE, 2021.

Zhang, Ji, Sanjiv Singh, and George Kantor. "Robust monocular visual odometry for a ground vehicle in undulating terrain." *Field and service robotics*. Springer, Berlin, Heidelberg, 2014.