Odometry Estimation Utilizing 6-DOF Force Sensors and IMU for Legged Robot

Huajian Wu
Department of Automation
Shanghai Jiao Tong University
Shanghai, China
summeryuki@sjtu.edu.cn

Yue Gao*
MoE Key Lab of Artificial Intelligence
Department of Automation
Shanghai Jiao Tong University
Shanghai, China
yuegao@sjtu.edu.cn

Shaoyuan Li
Department of Automation
Shanghai Jiao Tong University
Shanghai, China
syli@sjtu.edu.cn

2020 Chinese Automation Congress (CAC) | 978-1-7281-7687-1/20/$31.00 ©2020 IEEE | DOI: 10.1109/CAC51589.2020.9326974

***Abstract*—Odometry estimation is the problem of estimating platform pose utilizing different types of sensors. Visual odometry calculates the translation and rotation of two frames by extracting and matching feature points. However, it cannot function properly in some situations, such as dark environment or scenes with repeated textures. In rescuing scenario with poor visibility, fire-fighting robots and nuclear power plant rescue robots can not localize itself by solely depending on visual information. A real-time high precision odometry estimation method based on 6-DOF force sensors and IMU for legged robot, which can overcome the above limitations, is proposed in this paper. The odometry estimation method is realized by real-time robot posture estimation through 6-DOF force sensors mounted on the legs of legged robot. We combine particle filter and Quasi RNN network to build particle filter net. Through the fusion of 6-DOF force sensors and IMU using particle filter net, real-time odometry estimation with high precision is achieved. Compared with the state of art, our method has advantages of higher accuracy, faster computation speed and broader application scenarios for legged robot.**

***Index Terms*—odometry estimation, particle filter net, posture estimation, Quasi RNN network**

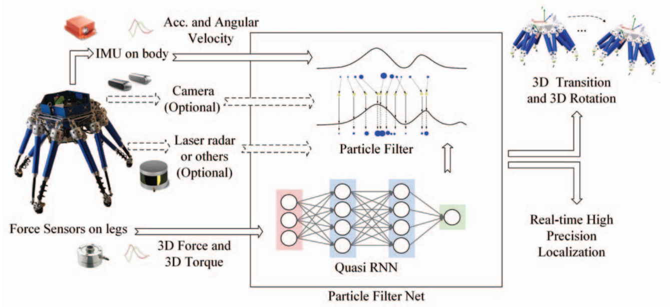

Fig. 1. This paper illustrates a real-time high precision odometry estimation method, which based on the 6-DOF force sensors and IMU for legged robots. Our method can work without vision. In our method, the Quasi RNN network is utilized to estimate the posture of the body according to the force and torque of legs. Then particle filter net is used to fuse force and torque information with IMU or other sensors to realize real-time high precision odometry estimation.

I. Introduction

Odometry estimation is the problem of estimating platform pose utilizing different types of sensors. The evaluation of odometry estimation is mainly focused on precision and real-time performance. A number of odometry estimation methods have been proposed recently [1] [2] [3] [4]. They are mostly based on vision or fusion of vision and IMU. Visual odometry calculates the displacement and rotation of two frames by extracting and matching feature points. However, it can not work in some situations, such as dark environment or scene with repeated textures. Robots are unable to locate accurately at night or in low-light indoor environment. In large warehouse, glass curtain wall or wood panel scene with repeated textures, vision sensors are difficult to extract and match feature points, hence high accuracy odometry estimation can not be achieved. In rescuing scenario with poor visibility, fire-fighting robots and nuclear power plant rescue robots can not localize itself by solely depending on visual information.

This work is sponsored by the National Natural Science Foundation of China (Grant No.61903247 and No.U1613208).

Odometry estimation can be achieved without vision sensors. Many researchers have proposed different sensor fusion methods for accurate odometry estimation without vision sensors. GPS and LiDAR can be utilized to obtain accurate position [5], but GPS does not work in indoor environment. To solve this problem, localization method using low cost MEMS-IMU sensors and terrain map in GPS denied environment is proposed [6]. However, MEMS sensors can not provide accurate localization estimation. On the other hand, pervasive magnetic field and opportunistic WiFi sensing can also be utilized to realize localization [7], but they are vulnerable to magnetic interference for outdoor environment.

In this paper, a real-time high precision odometry estimation method is proposed based on 6-DOF force sensors and IMU for legged robot. Our method can achieve good odometry estimation results without visual information. The proposed odometry estimation method is realized by robot posture estimation through fusing IMU and 6-DOF force sensors mounted on the legs of legged robot.

In particular, the main contributions of the paper include:

- Particle filter net method is proposed to fuse 6-DOF force sensors and IMU data with combination of deep learning and filter method. Our odometry estimation method has the advantages of high precision and real-time computation speed.

- Quasi RNN network is utilized to estimate the posture of the robot body according to the force and torque of legs. Compared with traditional methods, our method has great advantages of computation complexity and speed.
- Several experiments are conducted on six-legged robot in various indoor and outdoor terrain maps and the applicability of our method is shown.

II. Related Work

Visual odometry. ORB based odometry method is the state of art on visual odometry [8]. A large number of researchers have made a lot of improvements base on ORB visual odometry architecture [9] [10] [11]. Higher estimation accuracy can be obtained by using extended Kalman filter (EKF) to fuse vision and other sensors, such as IMU [12] [13] [14], LiDAR [15]. In nonlinear system, particle filter (PF) and improved particle filter, such as UPF and SSUPF performs better than extended Kalman filter (EKF). On the other hand, combination with points and line features [16] [17], feature extraction based on CNN [18] [19] and mixed data association [20] can also improve the estimation accuracy and real-time performance. Visual odometry extracts and matchs feature points between two frames to locate the robot. When video has a number of moving objects, it is difficult to locate these features accurately. Meanwhile, it can not work in scenes with repeated textures or environment with poor visibility.

Posture estimation for legged robot. Posture estimation solves the relationship between the 6-DOF force of legs and body posture, which is a nonlinear function. In tradition, with the 6-DOF force sensors, the exact pose of the legged robot can be computed using spiral theory [21] [22] and G_f set theory [22] [23] [24] . However, it can not meet the real-time application requirements because of its high computational complexity. Posture estimation is similar to the dynamic estimation. According to the previous researches, RNN network has a good approximation to robot dynamic model [25] [26]. RNN network, especially LSTM network containing gate control, has become the standard structures for deep learning to solve sequential tasks in recent years. RNN layer can not only solve the problem of variable length input, but also increase the depth of network and improve the representation ability through multi-layer stacking. However, the standard RNN (including the LSTM) is limited to handle very long sequences. Quasi RNN network combines the characteristics of RNN and CNN to improve parallelization and longseries capability [27]. In this paper, the Quasi RNN network is utilized to estimate the posture of the body.

III. Odometry Estimation Algorithm for Legged Robot

Real-time high precision odometry estimation method based on 6-DOF force sensors and IMU for legged robot is shown in Fig. 1. Vision sensor is an optional part in our method. At first, the Quasi RNN network is utilized to estimate the posture of the body according to the force and torque of legs. Then particle filter net is used to fuse 6-DOF force sensors information with IMU data to realize real-time high precision odometry estimation.

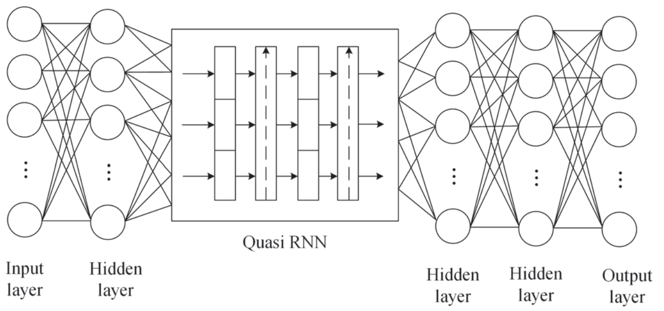

Fig. 2. The network structure. 6-DOF forces and torques of six legs are fed into Quasi RNN network, and output 128 dimensional features. Then they go through a three-layer full connection and the network output accurate posture, including 3D displacement and 3D Euler Angle. It can be divided into standing mode and walking mode according to the different number of input units.

A. Posture Estimation by Quasi RNN

In this paper, the Quasi RNN network is utilized to estimate the body posture according to the force and torque of the six legs, which achieves good accuracy and real-time performance.

The network structure used in this paper is shown in Fig. 2. 6-DOF force sensors data of six legs are fed into Quasi RNN network, and it outputs 128 dimensional features. Then they go through a three-layer full connection and the network outputs the accurate posture, including 3D displacement and 3D Euler Angle. It can be divided into standing mode and walking mode according to the different number of input units. When it is in standing mode, the number of input units is 36 and when it is in walking mode, the number of input units is 18. This network can use force and torque of legs to estimate the posture of the body.

1) Standing mode: When the robot stands on the ground, its six legs touch the ground steadily. According to the force and torque of six legs, the posture of the robot body can be calculated. The number of the input units is 36 and each leg contains a 3D force and a 3D torque. They are collected by the 6-DOF force sensors on each leg. The network outputs 6 dimensional posture, including 3D displacement and 3D Euler Angle.

2) Walking mode: When the robot walks by 3-3 gaits, three legs touch the ground steadily at each moment. According to the force and torque of three legs touched the ground, the posture of the robot body can be calculated. The number of input units is 18 and each leg contains a 3D force and a 3D torque. They are collected by 6-DOF force sensors on each leg. The network outputs 6 dimensional posture, including 3D displacement and 3D Euler Angle. If the robot walks by other gaits, the input of network should be changed to the legs which touching the ground and ignore the force information of legs off the ground.

Fig. 3. Experiment Scenes. (a) Small scene odometry estimation in dark environment. Visual odometry is not available when the robot is in the dark environment, but our method can also work well in this scene. (b) Large scene odometry estimation with repeated textures. Visual odometry is not available when the textures in the scene are repeated, but our method does not have this limit. (c) Posture estimation when standing. When the robot stands on the ground, it rotates its body and utilizes 6-DOF force sensors and IMU to estimate posture changes.

B. Particle Filter Net

The particle filter is based on monte carlo method, which expresses the distribution of random state using particles extracted from the posterior probability. In nonlinear system, it performs better than extended Kalman filter (EKF). For the nonlinear system which is difficult to build the exact model, particle filter net can get good filtering effect. Particle filter net uses neural network to approximate model and combines it with the traditional particle filter algorithm to filter the unknown states. This paper fuses the 6-DOF force sensors and IMU to estimate the robot's odometry. According to the $Newmark-\beta$ method, the transfer model of system is Eq. 1-5.

$$X_{k+1}^{body} = X_k^{body} + \dot{X}_k^{body}\Delta t + R_k^{body}\ddot{X}_k^{body}\frac{\Delta t^2}{2} \quad (1)$$

$$\dot{X}_{k+1}^{body} = \dot{X}_k^{body} + R_k^{body}\ddot{X}_k^{body}\Delta t \quad (2)$$

$$\ddot{X}_{k+1}^{body} = \ddot{X}_k^{body} \quad (3)$$

$$R_{k+1}^{body} = R_k^{body} Rot(W_k^{body}\Delta t) \quad (4)$$

$$W_{k+1}^{body} = W_k^{body} \quad (5)$$

X_k^{body}, $\dot{X}_k^{body}$ and $\ddot{X}_k^{body}$ are the representation of the world coordinates system (x, y, z) which are three dimensions. W_k^{body} is the angular velocity of the robot's body and R_k^{body} is the rotation matrix of robot body.

The observation model of each step is Eq. 6-8.

$$(X_k^{body}, R_k^{body}) = f(F_{leg_i}, T_{leg_i}, i = 1, ..., 3 \ or \ i = 1, ..., 6) \quad (6)$$

$$\ddot{X}_k^{body} = \ddot{X}_k^{body} \quad (7)$$

$$W_k^{body} = W_k^{body} \quad (8)$$

F_{leg_i}, T_{leg_i} are the force and torque of $leg_i (i = 1, ..., 3 \ or \ i = 1, ..., 6)$, which are collected by the 6-DOF force sensors on each legs. $\ddot{X}_k^{body}$ and W_k^{body} are measured by IMU. $f(F_{legi}, T_{legi}, i = 1, ..., 3 \ or \ i = 1, ..., 6)$ is the relationship between force and torque of legs and posture of body, which can be approximated by Quasi RNN network mentioned in Section III. When the robot stands on the ground, the 6-DOF force information of six legs are input into the Quasi RNN network and when the robot walks by 3-3 gait, the force information of three legs which touching the ground steadily are input into the network. Then the displacement and rotation of the robot body can be calculated by particle filter net for each step. Because the force and torque information is independent for each step, each step requires to restart a new particle filter to get real time location. Therefore, when the number of steps is large, this method has the problem of error accumulation, which will be discussed in the next section.

TABLE I
COMPARED WITH ORB-SLAM2

	Ours	ORB-SLAM2
Accuracy	0.0035 m/0.05°	0.035 m/0.25°
Speed	500 fps	30 fps
Scope	Robots with force sensors	Any robot with camera
Environment	Any environment	Limited environment

IV. Experiments

Experiments were conducted to verify whether the proposed method could accurately locate the robot, including small scene odometry estimation in dark environment, large scene odometry estimation with repeated textures and posture estimation when standing, as shown in Fig. 3.

The novel hexapod robot with P-P structure used in the experiments adopts the design of three-chain parallel structure, which greatly improves the bearing capacity, stiffness and precision of each single leg. It has the characteristics of strong topology. This allows the posture can be estimated accurately using 6-DOF force sensors mounted on six legs.

A. Compared With ORB-SLAM2

The comparison with ORB-SLAM2 is as shown in TABLE I. Compared with ORB-SLAM2, our method has higher accuracy and speed. The speed of our method mainly depends on the acquisition speed of force sensors and IMU. Our method can be utilized for robots with 6-DOF force sensors in any environment. ORB-SLAM2 is not work in some environments.

B. Network trainning

6-DOF force sensors are utilized to estimate the real-time body posture using Quasi RNN network, which is divided into standing and walking modes.

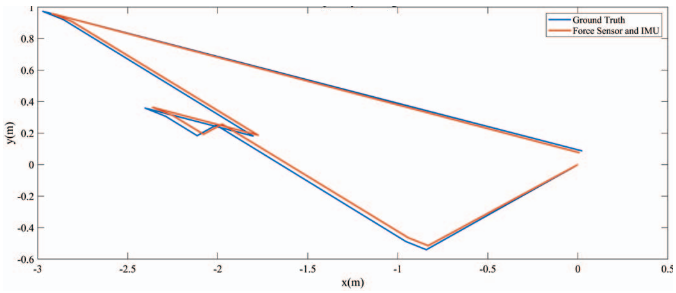

Fig. 4. Experiment result under world coordinate of small scene odometry estimation in dark environment. The blue line is ground truth and the red line represents the trajectory estimated by our method.

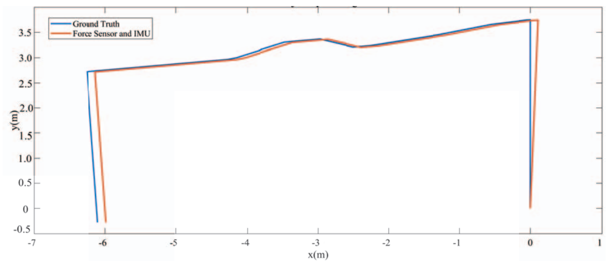

Fig. 5. Experiment result under world coordinate of large scene odometry estimation with repeated textures. The blue line is ground truth and the red line represents the trajectory estimated by our method.

1) Standing mode: The robot stands and moves the body in place, and then records the changes in 6-DOF force sensors and IMU. The ground truth can be obtained by high precision posture tracking camera, including 3D displacement and 3D Euler Angle. The number of the input units of the network is 36 and each leg contains a 3D force and a 3D torque. The training data covers the entire movement space of the robot body. The network training converges at about 50000 times, and the learning rate is 0.01. The accuracy of posture estimation is $0.005\ m\ 0.5°$.

2) Walking mode: The robot walks one step in different step length and direction by 3-3 gait, and then records the changes in 6-DOF force sensors and IMU. The ground truth can be obtained by high precision posture tracking camera, including displacement and Euler Angle. The number of input units of network is 18 and each leg, which touches the ground steadily, contains a 3D force and a 3D torque. The training data covers the entire movement space of one step. The network training converges at about 100000 times, and the learning rate is 0.01. The accuracy of posture estimation is $0.005\ m\ 0.5°$.

C. Small Scene Odometry Estimation in Dark Environment

Visual odometry is not available when the robot is in dark environment. The experiment is shown in Fig. 3(a). We make the robot move freely in small scene and record the trajectory. Then the 6-DOF force sensors and IMU are used to estimate the trajectory and compare the differences. The result is shown in Fig. 4. The coordinate is world coordinate. The blue line is ground truth and the red line represents the estimated trajectory by our method. Our method can estimate odometry accurately in small scenes dark environment.

D. Large Scene Odometry Estimation with Repeated Textures

Visual odometry estimation is not available when the textures in the scene are repeated as shown in Fig. 3(b). There are a lot of repeated textures in the scene such as glass and wood. Visual odometry is difficult to extract and match feature points but our method does not have this limit.

We make the robot move freely in large scene and record the trajectory. Then the 6-DOF force sensors and IMU are used to estimate the trajectory and compare the differences. The result is shown in Fig. 5. The coordinate is world coordinate. The blue line is ground truth and the red line represents the estimated trajectory by our method. Our method can locate accurately in large scenes with repeated texture but the error is larger than small scene because of error accumulation.

E. Posture Estimation When Standing

When the robot stands on the ground, it rotates its body and uses 6-DOF force sensors and IMU to estimate posture changes. The experiment is shown in Fig. 3(c).

The 6-DOF force sensors and IMU are used to estimate the trajectory and compare the differences. The result is shown in Fig. 6. The X axis is time, and the Y axis is Euler Angle. The blue line is ground truth and the red line represents the estimated trajectory by our method. The roll of IMU is measured by geomagnetism, which is inaccurate. Through the fusion of force sensor and IMU, the roll is measured more accurately.

V. Discussion and Conclusion

Odometry estimation based on 6-DOF force sensors and IMU for legged robot can overcome the limitations of visual odometry. The real-time odometry estimation of our method is realized by the fusion of 6-DOF force sensors and IMU using particle filter net. Compared with visual odometry like ORB-SLAM2, our method has advantages of higher accuracy and speed, and broader application scenes. When other sensors are available, it can combine with them to get higher accuracy easily.

Odometry estimation based on 6-DOF force sensors and IMU can be used in any robots with force sensors. Its applications scope extends far beyond legged robot. The framework of proposed method in this paper can also be used for snake-like robot, soft robot, etc. High DOF mechanical claw is another important application scope of our method. Through the haptic sensor array of the mechanical claw, real-time posture can be calculated. Particle filter net is used to fuse haptic information and other sensors such as IMU to get the real-time trajectory. It is of great significance for the high DOF mechanical claw to get the shape and texture of object by seeing and touching it using odometry estimation based on force sensors and IMU.

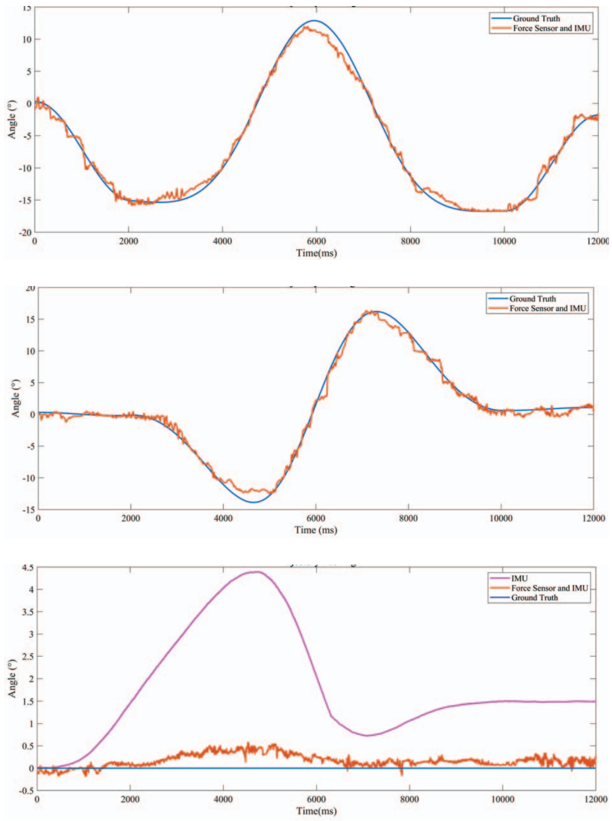

Fig. 6. Experiment result of odometry estimation when the standing posture changes. (a) the trajectory of α over time. (b) the trajectory of β over time. (c) the trajectory of γ compared with the measurement value of IMU. The blue line is ground truth and the red line represents the trajectory estimated by our method.

REFERENCES

[1] T. Qin, J. Pan, S. Cao, and S. Shen, "A general optimization-based framework for local odometry estimation with multiple sensors," *arXiv preprint arXiv:1901.03638*, 2019.

[2] B. Guan, P. Vasseur, C. Demonceaux, and F. Fraundorfer, "Visual odometry using a homography formulation with decoupled rotation and translation estimation using minimal solutions," in *2018 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2018, pp. 2320–2327.

[3] A. Lakshmi, A. Faheema, and D. Deodhare, "Robust direct visual odometry estimation for a monocular camera under rotations," *IEEE Robotics and Automation Letters*, vol. 3, no. 1, pp. 367–372, 2017.

[4] I. P. Alonso, D. F. Llorca, M. Gavilan, S. Á. Pardo, M. Á. García-Garrido, L. Vlacic, and M. Á. Sotelo, "Accurate global localization using visual odometry and digital maps on urban environments," *IEEE Transactions on Intelligent Transportation Systems*, vol. 13, no. 4, pp. 1535–1545, 2012.

[5] H. M. Cho, H. Jo, S. Lee, and E. Kim, "Odometry estimation via cnn using sparse lidar data," in *2019 16th International Conference on Ubiquitous Robots (UR)*. IEEE, 2019, pp. 124–127.

[6] H. Ahmed and M. Tahir, "Terrain-based vehicle localization using low cost mems-imu sensors," in *2016 IEEE 83rd Vehicular Technology Conference (VTC Spring)*. IEEE, 2016, pp. 1–5.

[7] Y. Shu, C. Bo, G. Shen, C. Zhao, L. Li, and F. Zhao, "Magicol: Indoor localization using pervasive magnetic field and opportunistic wifi sensing," *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 7, pp. 1443–1457, 2015.

[8] R. Mur-Artal and J. D. Tardós, "Orb-slam2: An open-source slam system for monocular, stereo, and rgb-d cameras," *IEEE Transactions on Robotics*, vol. 33, no. 5, pp. 1255–1262, 2017.

[9] P. Burdziakowski, "Towards precise visual navigation and direct georeferencing for mav using orb-slam2," in *2017 Baltic Geodetic Congress (BGC Geomatics)*. IEEE, 2017, pp. 394–398.

[10] B. Bescos, J. M. Fácil, J. Civera, and J. Neira, "Dynaslam: Tracking, mapping, and inpainting in dynamic scenes," *IEEE Robotics and Automation Letters*, vol. 3, no. 4, pp. 4076–4083, 2018.

[11] X. Zuo, X. Xie, Y. Liu, and G. Huang, "Robust visual slam with point and line features," in *2017 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2017, pp. 1775–1782.

[12] G. Nützi, S. Weiss, D. Scaramuzza, and R. Siegwart, "Fusion of imu and vision for absolute scale estimation in monocular slam," *Journal of intelligent & robotic systems*, vol. 61, no. 1-4, pp. 287–299, 2011.

[13] M. W. Achtelik, S. Lynen, S. Weiss, L. Kneip, M. Chli, and R. Siegwart, "Visual-inertial slam for a small helicopter in large outdoor environments," in *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*. IEEE, 2012, pp. 2651–2652.

[14] P. Geneva, J. Maley, and G. Huang, "An efficient schmidt-ekf for 3d visual-inertial slam," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019, pp. 12 105–12 115.

[15] S.-H. Chan, P.-T. Wu, and L.-C. Fu, "Robust 2d indoor localization through laser slam and visual slam fusion," in *2018 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, 2018, pp. 1263–1268.

[16] X. Wei, J. Huang, and X. Ma, "Real-time monocular visual slam by combining points and lines," in *2019 IEEE International Conference on Multimedia and Expo (ICME)*. IEEE, 2019, pp. 103–108.

[17] D. Wu, Y. Meng, K. Zhan, and F. Ma, "A lidar slam based on point-line features for underground mining vehicle," in *2018 Chinese Automation Congress (CAC)*. IEEE, pp. 2879–2883.

[18] P. Maolanon, K. Sukvichai, N. Chayopitak, and A. Takahashi, "Indoor room identify and mapping with virtual based slam using furnitures and household objects relationship based on cnns," in *2019 10th International Conference of Information and Communication Technology for Embedded Systems (IC-ICTES)*. IEEE, 2019, pp. 1–6.

[19] P. Tang, Z. Wang, N. Qi, and Q. Zhu, "A fast feature extraction process for visual slam," in *2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*. IEEE, 2019, pp. 959–963.

[20] F. Wen, X. Chai, Y. Li, W. Zou, K. Yuan, and P. Chen, "An improved visual slam algorithm based on mixed data association," in *2011 9th World Congress on Intelligent Control and Automation*. IEEE, 2011, pp. 1081–1086.

[21] S. Ma, "Analysis of snake movement forms for realization of snake-like robots," in *Proceedings 1999 IEEE International Conference on Robotics and Automation (Cat. No. 99CH36288C)*, vol. 4. IEEE, 1999, pp. 3007–3013.

[22] J. Yang, F. Gao, Z. Jin, and L. Shi, "Classification of lying states for the humanoid robot sjtu-hr1," *Science in China Series E: Technological Sciences*, vol. 52, no. 5, pp. 1301–1311, 2009.

[23] J. Yang and F. Gao, "Classification of sitting states for the humanoid robot sjtu-hr1," *Journal of Bionic Engineering*, vol. 8, no. 1, pp. 49–55, 2011.

[24] J. He, F. Gao, X. Meng, and W. Guo, "Type synthesis for 4-dof parallel press mechanism using g f set theory," *Chinese Journal of Mechanical Engineering*, vol. 28, no. 4, pp. 851–859, 2015.

[25] M. Oubbati, M. Schanz, and P. Levi, "Kinematic and dynamic adaptive control of a nonholonomic mobile robot using a rnn," in *2005 International Symposium on Computational Intelligence in Robotics and Automation*. IEEE, 2005, pp. 27–33.

[26] H. Shao, K. Nonami, T. Wojtara, R. Yuasa, S. Amano, and D. Waterman, "Neuro-fuzzy position control of demining tele-operation system based on rnn modeling," *Robotics and Computer-Integrated Manufacturing*, vol. 22, no. 1, pp. 25–32, 2006.

[27] J. Bradbury, S. Merity, C. Xiong, and R. Socher, "Quasi-recurrent neural networks," *arXiv preprint arXiv:1611.01576*, 2016.