Low-light Image Enhancement for Construction Robot Simultaneous Localization and Mapping

Xinyu Chen[1] and Yantao Yu[1]

[1]Department of Civil and Environmental Engineering, Hong Kong University of Science and Technology

xchengl@connect.ust.hk, ceyantao@ust.hk

Abstract -

Visual Simultaneous Localization and Mapping (V-SLAM) is widely used in construction robots because it is an efficient and inexpensive information acquisition method. However, low-light construction scenes pose significant challenges for V-SLAM detection and positioning. In low-light scenes such as underground garages or dim indoor scenes, V-SLAM is difficult to detect enough valid feature points, which causes navigation to fail. To address this issue, we propose an Unsupervised V-SLAM Light Enhancement Network (UVLE-Net) to enhance low-light images. After image enhancement, we add a robust Shi-Tomasi method in ORB-SLAM2 to detect feature points and use the sparse optical flow algorithm to track the feature points. By using UVLE-Net, the brightness and contrast of the images can be significantly increased, and feature points can be detected easily. The optical flow and Shi-Tomasi method improve the ability of feature point extraction and tracking in low light. To validate the robustness and superiority of our method in low-light conditions, we conduct comparison experiments with other enhancement techniques on published and real-world construction datasets.

Keywords -

Low-light enhancement; Retinex; Simultaneous Localization and Mapping; Construction robots

1 Introduction

The productivity, quality, and other aspects of construction have significantly increased as a result of the rapid development of construction robotics[1]. Autonomous mobile robots can perform specific tasks by navigating the construction site[2]. For example, a waste recycling robot can inspect construction sites and find nails and screws[3]. Construction monitoring robots with scanning sensors and cameras can move automatically and monitor progress [4]. The autonomous task of construction robots depends on the navigation system. Mobile robots typically use Simultaneous Localization and Mapping (SLAM) technology to generate maps of unknown environments while locating robots in the map. Construction robots can move and work automatically using the SLAM system.

Construction robots need to collect enough on-site data in order to operate autonomously. At present, laser point clouds and RGB cameras are mainly used in positioning and working system. Visual data contains rich data information, and provides a data basis for construction site work. In comparison to other SLAM sensors like LiDAR, the RGB camera is typically lightweight, inexpensive, and contains rich visual information. State-of-the-art visual-based algorithms such as ORB-SLAM2 perform well in both datasets and real-world experiments[5]. However, the SLAM algorithm is set in a well-lit environment[6][7]. The low-light construction sites pose significant challenges for V-SLAM. For example, in an interior building environment, a robot moving from an area with good lighting to an area far away from artificial lighting facilities, such as a backlit area or an underground garage with many occlusions, can experience location errors or even failure. Insufficient light in image shooting can significantly reduce the visibility of the image. V-SLAM methods rely on feature point extraction, but low-light frames provide few high-quality points. Therefore, V-SLAM systems would lose their localization and detection abilities in low light conditions due to low light and low contrast. Moreover, due to dynamic lighting and halos in a dark scene, it is difficult to track the same feature points as the lighting changes. In order to improve the visibility of the low-light construction environment and the usability of the current robot navigation and positioning system, low-light images need to be enhanced.

Although there have been some studies on enhancing low-light images in the field of computer vision, most low-light image enhancement techniques, such as LIME[8], Low-light GAN[9], MSRCR[10], are optimized to enhance the image perception by human eyes rather than the V-SLAM performance. The existing low-light image enhancement approaches have limited ability to address low-light difficulties in construction robots V-SLAM, given the following existing problems: (a) Due to the limitation of manual parameter setting, the manual enhancement method based on image features can only solve one of the problems in challenging low-light construction scenarios. (b) Data-driven enhancement methods require training on pairs of data sets. In dynamic construction environments, it is challenging to collect paired data sets with exactly the same content. (c) The existing low-light enhancement methods lack research on feature point extraction of construction robots V-SLAM. Therefore, it is necessary to build a low-light enhanced method for construction robot localization and mapping in low light.

In this paper, we propose an Unsupervised V-SLAM Light Enhancement Network (UVLE-Net) to enhance low-light images, which helps construction robots to locate under low light. To address the issue of construction robots in low light, we preprocess the image using the deep neural network of the Retinex theor y. Following image preprocessing, we detect feature points with the Shi-Tomasi method and uniformly assign feature points to track. Finally, we test UVLE-Net on the EuRoc dataset[11] and construction environment. Experiments show that our proposed method has better detection performance in both open datasets and real low-light construction environments. The contributions of our work are summarized as follows: (a) We propose an UVLE-Net for low-light construction robots to enhance frames' lightness and contrast. It can help construction robots to get more gradient information and feature points in low light. (b) We design novel loss functions to train the image enhancement network in an unsupervised way, thereby drastically reducing the requirement of coupled image collection. (c) We combine the Shi-Tomasi and sparse optical flow algorithm into ORB-SLAM2 to detect and track the feature points in low-light construction. It can usefully reduce the requirement of ORB-SLAM2 feature points and improve positioning accuracy in low-texture and low-light construction environments.

Our approach is innovative in three aspects: (a) Based on Retinex theory, we use a deep neural network to iteratively solve the reflectance map of the image, which has better generalization performance in construction scenes. (b) The enhancement network can be trained on any lightness dataset by designing unsupervised learning mode loss functions. (c) The Shi-Tomasi and sparse optical flow method is introduced in ORB-SLAM2 to increase its low-light positioning ability.

2 Literature review

2.1 SLAM in construction robotics

In construction, SLAM has gained more and more attention from researchers as an effective autonomous movement and location method[12]. SLAM algorithms can be divided into LiDAR-based and visual-based methods. Regarding LiDAR SLAM, Cebollada et al.[13] used a monocular camera with a 2-D laser sensor to locate and map the underground space. It used the bilinear filter function to estimate grid occupancy and the Gauss-Newton method to optimize scan matching. However, these methods lack closed-loop detection and optimization. Therefore, it is not capable of eliminating accumulated localization errors. Kim et al.[14] used 3-D laser SLAM to locate and match point cloud objects. It is limited by laser detection range and cannot be modelled in large construction scenes. It costs lots of computing resources in point cloud location and matching. Meanwhile, laser SLAM lacks the semantic information of the maps, which cannot help robots to do some work.

Regarding visual SLAM, it only needs inexpensive cameras and a few computing resources. It is economical and suitable for construction environments. Peel et al.[15] combined adaptive monte carlo localization with SLAM, using a small robot to detect bridge supports. Zhang et al.[16] used mobile robots for large-scale 3D printing construction. These robots mapped the environment using Gmapping and adaptive monte carlo localization to locate themselves. Asadi et al.[17] proposed a mobile robot platform equipped with an RGB camera. It used visual SLAM and semantic segmentation techniques to navigate daytime construction scenes. Although there has been a lot of work to demonstrate the effectiveness of SLAM in construction, most of them have been done under sufficient lighting conditions. These methods will not work in low-light conditions. In the low-light and low-texture construction environment, visual SLAM hardly extracts and tracks feature points due to the loss of gradient information. So there are some robustness and reliability issues with V-SLAM in a low-light environment.

2.2 Low-light image enhancement

The low-light image enhancement methods can be divided into conventional and data-driven methods. Regarding conventional methods, local statistics and intensity mapping are the primary earlier conventional methods. Histogram-based solutions expand the enhanced dynamic range by modifying the light distribution of global or local images[18]. Then, researchers used the Retinex theory to enhance images. Reflection mapping was used as the result of enhancement in Multiscale Retinex with Color Restoration (MSRCR) [10]. To enhance image lighting conditions, Fu et al.[19] used a weighted variational model to estimate the reflectivity and illuminance of images. By determining the RGB channel's maximum pixel intensity for each pixel, LIME[8] calculated the rough illumination map. Most conventional methods depend on carefully designed parameters and fail to generalize to various low-light construction conditions.

With the development of deep neural networks, many enhanced methods use deep learning techniques. For instance, the first low-light enhanced network in the GAN model was low-lightgan[9]. To simulate lighting images, it created a discriminator and generator using paired

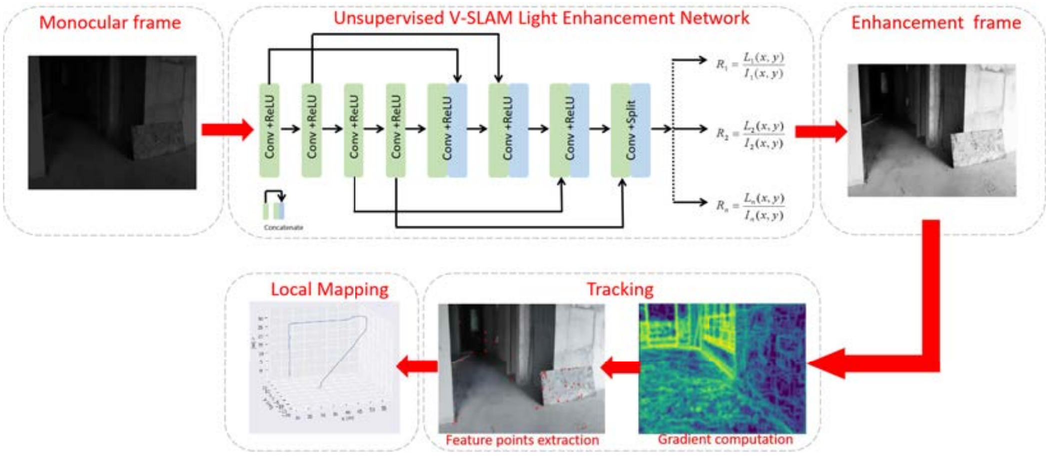

Figure 1. Architecture of the low-light construction robot navigation system.

Table 1. Detailed architecture of UVLE-Net.

Input	Dimensions	Operator	Output
Image	$256 \times 256 \times 1$	Conv+ReLU	Conv1
Conv1	$256 \times 256 \times 32$	Conv+ReLU	Conv2
Conv2	$256 \times 256 \times 32$	Conv+ReLU	Conv3
Conv3	$256 \times 256 \times 32$	Conv+ReLU	Conv4
Conv1+Conv4	$256 \times 256 \times 64$	Conv+ReLU	Conv5
Conv2+Conv5	$256 \times 256 \times 64$	Conv+ReLU	Conv6
Conv3+Conv6	$256 \times 256 \times 64$	Conv+ReLU	Conv7
Conv4+Conv7	$256 \times 256 \times 64$	Conv+ReLU	Conv8
Conv7+Conv8	$256 \times 256 \times 8$	Split	$R_1 \cdots R_n$
Image$\times R_n$	-	-	Enhanced image

datasets. Zero-DCE [20] extended the solution by establishing numerous particular image curves based on zero-inference. However, the model based on light recovery curves exhibits unstable performance under different levels of darkness. Therefore, in order to enhance the robustness of the image, it is necessary to estimate the image features.

3 Methodology

In this section, we first introduce the Retinex theory. On this basis, we present an Unsupervised V-SLAM Light Enhancement Network (UVLE-Net). Then, we introduce our robot positioning mapping method based on ORB-SLAM2. Figure 1 shows the overall structure of our method. First, the low-light images are fed into the enhancement network. We incorporate the Retinex theory into a deep neural network model to solve the reflectance iteratively. Our data-driven reflectance model can enhance images accurately under various illumination conditions. Then, ORB-SLAM2 computes gradients and feature points for robot localization and mapping.

3.1 UVLE-Net for low-light image enhancement

The traditional Retinex theory simulates human color perception. It is assumed that the image can be broken down into two parts: reflectance and illumination. The image $I(x,y) \in R^{W \times H \times 1}$ represents the source image, then it can be decomposed by

$$I(x,y) = R(x,y) \cdot L(x,y), \tag{1}$$

where $L(x,y)$ is the spatial distribution of source illumination, $R(x,y)$ denotes the distribution of scene reflectance. Reflectance denotes the intrinsic property of captured objects, which is consistent under all illumination conditions. The illumination depicts the varying degrees of brightness of objects.

We attempt to estimate the reflectance as guidance for automatic low-light image enhancement, with the merits of a simple and differentiable expression relying on the input images and preserving the differences of neighbouring pixels. Assuming that the illumination map under normal light condition L is an identity matrix, according to the effective formulation of Retinex, the reflectance for enhancement can be obtained. Arguably, the formulation of $R(x,y)$ is an ill-posed problem, and direct decomposition results in unnatural artifacts. So we design the enhanced model and use iterative algorithms to gradually eliminate the impact, which is

$$\log R_i(x, y) = \log I_i(x, y) - \log L_i(x, y), \quad \forall i = 1 \dots n. \tag{2}$$

where n is the number of iterations. Here we set $n = 5$ empirically, which we will detailedly discuss in the ablation study.

The detailed architecture of the enhanced model is demonstrated in TABLE 1. Input frames in SLAM are grey images, so the UVLE-Net is designed to enhance brightness in grey images. The proposed enhanced images model consists of 8 convolutional layers with skip connections. Specifically, the first 7 layers have 32 convolutional kernels of size 3×3 and stride 1 followed by the ReLU activation function, and the last one has 8 convolutional kernels of size 3×3 and stride 1 followed by the Tanh activation function. Essentially, the last convolutional layer splits the estimated lighting reflectance R_i, and the given image I_i is enhanced iteratively in terms of the parameter maps R_1 to R_n.

It is challenging to capture different lighting-paired images of a uniform dynamic construction scene. So we adopt unsupervised learning to overcome this problem. We devise several differentiable loss functions to facilitate the unsupervised reflectance illumination model training.

Loss functions:

(1) Information Difference Loss. In robot location tasks, the quality of feature point extraction is important. Thereby, an information difference loss is devised to quantify the differences between the improved image and the original image, which is expressed as:

$$L_{idl} = \| V(I_n) - V(I_0) \|_2^2, \tag{3}$$

where $V(I)$ is the feature extraction operator based on VGG-16. I_0 is the original raw image and I_n is the corresponding enhanced image after n time iterations. The VGG network is leveraged here for its concise architecture to compute information differences effectively.

(2) Exposure Control Loss. The stable exposure intensity in SLAM is the key to the position. Therefore, the exposure control loss is required to equalize exposure. To obtain an average intensity Y, the image is split into $16 * 16$ non-overlapping local regions. According to exposure fusion theory, a well-exposedness level E is defined as the grey level in grey space. As a result, the exposure control loss, which measures the distance between average intensity Y and well-exposedness level E, is calculated as follows:

$$L_{ecl} = \frac{1}{M} \sum_{m=1}^{M} |Y_m - E|, \tag{4}$$

where M represents the number of non-overlapping local regions, E is set as 0.6 empirically.

(3) Illumination Smoothness Loss. An illumination smoothness loss is applied to the estimated reflectance parameter map Ri to ensure the smoothness of the illumination component in grey space for maintaining the monotonicity of pixel-level surrounding context during iteration, which is expressed as:

$$L_{isl} = \frac{1}{N} \sum_{i=1}^{N} \left(|\nabla_x R_i| + |\nabla_y R_i| \right)^2, \tag{5}$$

where N represents the number of iterations, ∇ is the first-order differential operator, accordingly ∇_x and ∇_y denote the horizontal and vertical gradient operations respectively. We will explain them in the experimental section.

To sum up, the total loss is expressed as:

$$L = \omega_a L_{idl} + \omega_b L_{ecl} + \omega_c L_{isl}, \tag{6}$$

where ω_a, ω_b, and ω_c are the weights controlling the significance of losses.

3.2 ORB-SLAM2 for localization and mapping

In this paper, ORB-SLAM2 is selected as the navigation algorithm for construction robots. The workflow of ORB-SLAM2 consists of four steps: detection of feature points, stereo matching, feature tracking, and motion estimation. When ORB-SLAM2 receives a frame, it first extracts the feature points. A stereo-matching process follows feature detection. Then the field points created in the previous frame are projected to the current frame and matched with the current frame to obtain several corresponding features. The ORB-SLAM2 system relies on extracting and matching many feature points, and all the foundations are built on the accurate extraction of rich feature points. Therefore, system position and pose tracking will fail when operating in a construction scene with low texture and low light.

We adopt the Shi-Tomasi[21] and Kanade-Lucas-Tomasi (KLT)[22] sparse optical flow algorithm to solve these problems. The average distribution of feature points in the low illumination environment reduces the matching requirement of feature points. New feature points will be added if not enough features are left after the uniform procedure to reach the necessary 200. The next frame will be considered a new keyframe if the difference between the two frames' parallaxes is significant.

4 Experiments and results

The main focus of our work is the visual navigation of construction robots to work reliably in low-light environments. Therefore, we experiment on low-light open datasets and real construction scenes. First, we train the UVLE-Net model. Then we compare our method with other low-light enhancement methods and evaluate their accuracy on public datasets. Finally, we test them in real low-light construction scenes.

4.1 UVLE-Net training and ablation experiment

The UVLE-Net is trained in an unsupervised way on SICE[23]. SICE is a dataset consisting of 589 image sequences and 4,413 high-resolution samples with significant exposure levels. We use multi-exposure datasets to train the enhanced image model. It can reduce the impact of varying lighting conditions. These datasets are guaranteed to improve the generalization performance of UVLE-Net under various lighting conditions in construction scenes. UVLE-Net training experiments are conducted on a desktop Dell 7000 workstation. It has an Intel I7-12700 CPU running at 3 GHz, 16 GB of RAM, and an NVIDIA 3060 GPU. The network is optimized using the ADAM optimizer with default parameters and a fixed learning rate of 0.001. The weights of loss functions: w_a,w_b, and w_c are set as 0.01, 50, and 100 empirically.

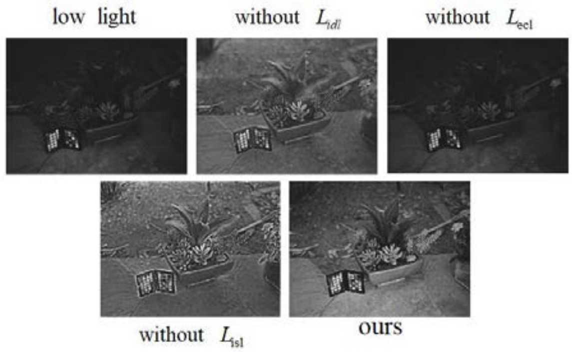

Figure 2. Ablation study on the contribution of loss functions.

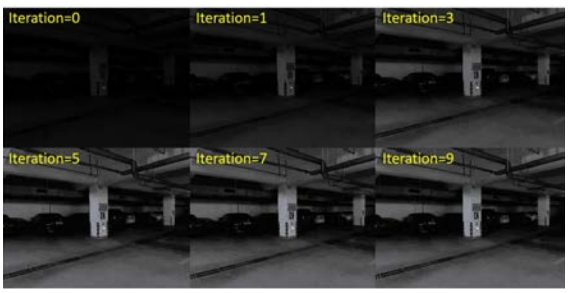

Figure 3. Qualitative comparison of the different iterations.

We use the same dataset and parameter to train the UVLE-Net model with different loss functions to verify the effects of each loss function, as shown in Figure 2. The loss of information difference L_{idl} enhanced the features produced from the input image, improving the model's ability to interpret semantics. Without the loss of exposure control L_{ecl}, the brightness decreases, and the low-light area fails to recover. The illumination smoothness loss L_{isl} acts as a link between surrounding pixel-level changes, Ensuring the unity of the overall brightness. We respectively test n = 0, 1, 3, 5, 7 and 9 iterations. It reveals that the peak performance appears when n equals five. Figure 3 depicts a visual example adapted to the different number of iterations. Increasing the number of iterations improves the lightness intuitively. However, when n > 5, performance barely improves. To verify the availability of our method in low light, we compare it with other advanced enhancement methods in SLAM feature point extraction and gradient calculation, as shown in Figure 4. It can be seen that the global overexposure of MSRCR reduces the information gradient and makes it difficult to extract feature points. The LIME approach has local confusion and ambiguity. Our method can improve brightness and contrast steadily. It has clearer gradient information and more feature points.

4.2 Low-light SLAM experiment

All experiments are conducted on a Ubuntu 22.04 laptop with an Intel I7-7700 CPU at 2.1Hz and ROS kinetic energy 2GB of memory. Camera FPS in the EuRoc data and real construction data are 20Hz, and UVLE-Net can run at 22Hz on CPU and 67Hz on GPU. So it can run in real time.

A.Public dataset experiment

We use MH sequences in EuRoc MAV as the test datasets. The EuRoc MAV datasets consist of 5 sequence micro aircraft vehicles (MAVs) flying rooms around two different directions (V1 and V2 sequences) and one large industrial (MH sequences). Depending on MAVs' speed, lighting, and texture, the sequences are classified as easy, medium, and difficult. In MH datasets, there are many low-light and complex light scenes. These indoor lighting problems are very similar to low-light problems in construction scenes, so they are used to test algorithms in low-light environments. As shown in Figure 5, we use MSRCR, LIME and UVLE-Net to enhance low-light images in MH datasets.

Table 2 compares the performance of UVLE-Net in monocular sensors with original, MSRCR and LIME; all tests are in KLT ORB-SLAM2. As shown in the table, our method achieves lower error results in all sequences than other enhanced methods, in most cases by a large margin. In most sequences, using enhancement methods can reduce the error. Nevertheless, MSRCR produces a larger error because overexposure destroys gradient and

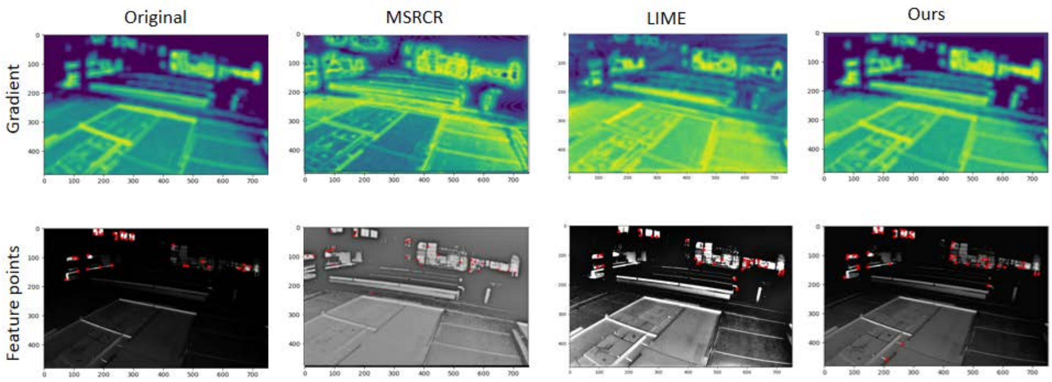

Figure 4. The gradient and feature point results of original, MSRCR, LIME and our method.

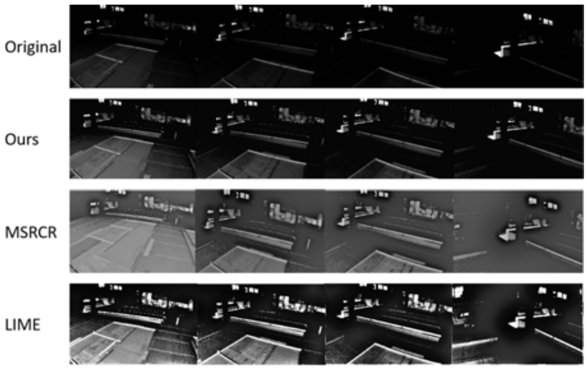

Figure 5. The results of image preprocessing in MH datasets. MSRCR shows global overexposure, LIME has local overexposure and blurring, and our method steadily increases brightness, showing more detail and no overexposure.

Table 2. Performance comparison in the EuRoC MH datasets (Root mean square error in m).

Sequence	Original	MSRCR	LIME	Ours
MH01 easy	0.039	0.031	0.034	**0.024**
MH02 easy	0.036	0.028	0.033	**0.025**
MH03 medium	0.055	0.059	0.055	**0.053**
MH04 difficult	0.089	0.084	0.079	**0.073**
MH05 difficult	0.077	0.062	0.064	**0.055**

feature points in some normally lit scenes in MH-03. In addition to the table, we demonstrate the result of the MH-01 sequence on both 2D and 3D pose graphs in Figure 6 and Figure 7. In the pose graphs map, X-Z represents the ground, and Y represents the height. From the enlarged part in the figures, we can see that under the original low-light environment, the SLAM algorithm has a significant error in this section. MSRCR and LIME can reduce these errors. The estimated trajectory of our method is almost consistent with the real value.

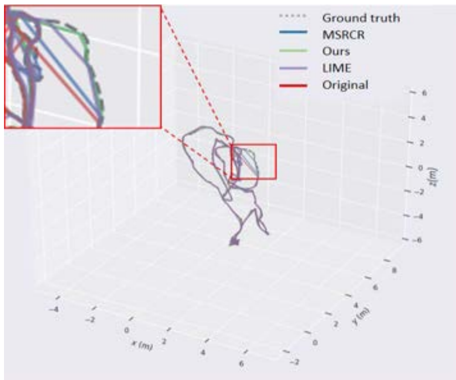

Figure 6. 3D pose graph. The broken line is SLAM ground truth. The blue line is MSRCR. The green line is our method. The purple line is LIME. The red line is the original frames(X-Z represents the ground, and Y represents the height).

B.Real construction dataset experiment

We use the Intel Realsense 435 camera to conduct real experiments in low-light indoor construction scenes. We lack the pose motion capture system in construction scenes to collect the actual moving position. So we only compare our low-light enhanced method with the original method in the real-world experiment. In Figure 8, we compare the original image with our enhanced image. The processed image has higher brightness and contrast. It shows more gradient information and feature points. In order to further verify the performance of our algorithm, we conduct location estimation and mapping experiments. The results are shown in Figure9. As can be seen, in the latter part of the navigation (left part), the original frames cannot provide enough feature points for tracking and matching, so the mapping is lost. In addition, there is a large deviation in the corner part.

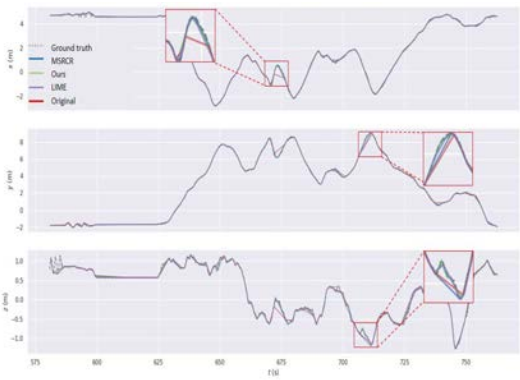

Figure 7. 2D pose graph. The broken line is SLAM ground truth. The blue line is MSRCR. The green line is our method. The purple line is LIME. The red line is the original frames.

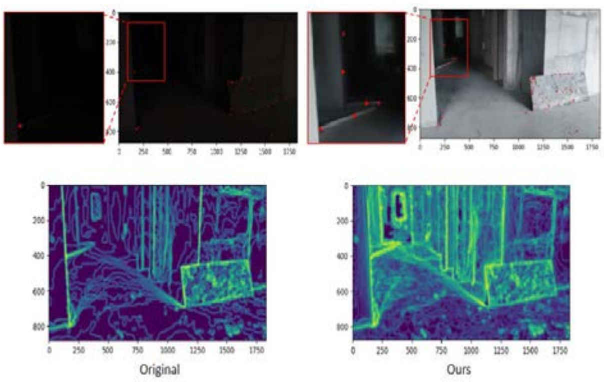

Figure 8. Original image captured by the Intel Realsense 435 camera and processed image in UVLE-Net.

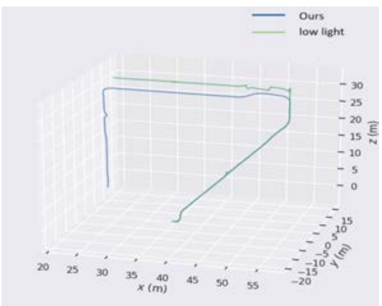

Figure 9. 3D pose graph. The green line is the original method in low light. The blue line is our method(X-Z represents the ground, and Y represents the height).

5 Conclusions

This paper proposes a visual-based monocular construction robot SLAM system for low-light challenges. We design a low-light enhancement network for robot construction scenes. The deep network based on Retinex theory trains the enhancement model using self-supervised learning. This method improves the brightness and contrast of the frame and helps the construction robot to track feature points in low light environments. We add the Shi-Tomasi and optical flow method to the ORB-SLAM2 system to reduce the difficulty of feature point tracking in the SLAM system. We validate our method in the public and real-world datasets. The results show that our method has better performance than MSRCR and LIME. Our method solves the positioning failure problem in low-light construction scenes. Our method can be run in real-time on mobile computers and it is suitable for deploying robots or portable devices in low-light environments.

6 Acknowledgement

This work was financially supported by the HKUST-BDR Joint Research Institute Fund, No. OKT23EG01.

References

[1] Xinge Zhao and Chien Chern Cheah. Bim-based indoor mobile robot initialization for construction automation using object detection. *Automation in Construction*, 146:104647, 2023. ISSN 0926-5805.

[2] Hongling Guo, Jia-Rui Lin, and Yantao Yu. Intelligent and computer technologies' application in construction, 2023.

[3] Zeli Wang, Heng Li, and Xiaoling Zhang. Construction waste recycling robot for nails and screws: Computer vision technology and neural network approach. *Automation in Construction*, 97:220–228, 2019.

[4] Muhammad Ilyas, Hui Ying Khaw, Nithish Muthuchamy Selvaraj, Yuxin Jin, Xinge Zhao, and Chien Chern Cheah. Robot-assisted object detection for construction automation: data and information-driven approach. *IEEE/Asme Transactions on Mechatronics*, 26(6):2845–2856, 2021.

[5] Raul Mur-Artal and Juan D Tardós. Orb-slam2: An open-source slam system for monocular, stereo, and rgb-d cameras. *IEEE transactions on robotics*, 33 (5):1255–1262, 2017.

[6] Myriam Servières, Valérie Renaudin, Alexis Dupuis, and Nicolas Antigny. Visual and visual-inertial slam: State of the art, classification, and experimental benchmarking. *Journal of Sensors*, 2021, 2021.

[7] Xinyu Chen and Yantao Yu. Image illumination enhancement for construction worker pose estimation in low-light conditions. In *Computer Vision–ECCV 2022 Workshops: Tel Aviv, Israel, October 23–27, 2022, Proceedings, Part VII*, pages 147–162. Springer, 2023.

[8] Xiaojie Guo, Yu Li, and Haibin Ling. Lime: Low-light image enhancement via illumination map estimation. *IEEE Transactions on image processing*, 26 (2):982–993, 2016.

[9] Guisik Kim, Dokyeong Kwon, and Junseok Kwon. Low-lightgan: Low-light enhancement via advanced generative adversarial network with task-driven training. In *2019 IEEE International Conference on Image Processing (ICIP)*, pages 2811–2815. IEEE, 2019.

[10] Daniel J Jobson, Zia-ur Rahman, and Glenn A Woodell. A multiscale retinex for bridging the gap between color images and the human observation of scenes. *IEEE Transactions on Image processing*, 6(7):965–976, 1997.

[11] Michael Burri, Janosch Nikolic, Pascal Gohl, Thomas Schneider, Joern Rehder, Sammy Omari, Markus W Achtelik, and Roland Siegwart. The euroc micro aerial vehicle datasets. *The International Journal of Robotics Research*, 35(10):1157–1163, 2016.

[12] Lichao Xu, Chen Feng, Vineet R Kamat, and Carol C Menassa. An occupancy grid mapping enhanced visual slam for real-time locating applications in indoor gps-denied environments. *Automation in Construction*, 104:230–245, 2019.

[13] Sergio Cebollada, Luis Paya, Miguel Julia, Mathew Holloway, and Oscar Reinoso. Mapping and localization module in a mobile robot for insulating building crawl spaces. *Automation in Construction*, 87: 248–262, 2018.

[14] Pileun Kim, Jingdao Chen, and Yong K Cho. Slam-driven robotic mapping and registration of 3d point clouds. *Automation in Construction*, 89:38–48, 2018.

[15] H Peel, S Luo, AG Cohn, and R Fuentes. Localisation of a mobile robot for bridge bearing inspection. *Automation in Construction*, 94:244–256, 2018.

[16] Xu Zhang, Mingyang Li, Jian Hui Lim, Yiwei Weng, Yi Wei Daniel Tay, Hung Pham, and Quang-Cuong Pham. Large-scale 3d printing by a team of mobile robots. *Automation in Construction*, 95:98–106, 2018.

[17] Khashayar Asadi, Hariharan Ramshankar, Harish Pullagurla, Aishwarya Bhandare, Suraj Shanbhag, Pooja Mehta, Spondon Kundu, Kevin Han, Edgar Lobaton, and Tianfu Wu. Vision-based integrated mobile robotic system for real-time applications in construction. *Automation in Construction*, 96:470–482, 2018.

[18] Chulwoo Lee, Chul Lee, and Chang-Su Kim. Contrast enhancement based on layered difference representation of 2d histograms. *IEEE transactions on image processing*, 22(12):5372–5384, 2013.

[19] Ying Fu, Antony Lam, Imari Sato, Takahiro Okabe, and Yoichi Sato. Separating reflective and fluorescent components using high frequency illumination in the spectral domain. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 457–464, 2013.

[20] Chunle Guo, Chongyi Li, Jichang Guo, Chen Change Loy, Junhui Hou, Sam Kwong, and Runmin Cong. Zero-reference deep curve estimation for low-light image enhancement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1780–1789, 2020.

[21] Jianbo Shi et al. Tomasi. good features to track. In *1994 Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, pages 593–600. sn, 1994.

[22] Myung Hwangbo, Jun-Sik Kim, and Takeo Kanade. Inertial-aided klt feature tracking for a moving camera. In *2009 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 1909–1916. IEEE, 2009.

[23] Jianrui Cai, Shuhang Gu, and Lei Zhang. Learning a deep single image contrast enhancer from multi-exposure images. *IEEE Transactions on Image Processing*, 27(4):2049–2062, 2018.