

# Passive Infrared Image Detail Enhancement with Points and Lines for Visual Odometry

Fan Li<sup>a</sup>, Xuezhao Zhang<sup>b</sup>, Yaohui Zhang<sup>c</sup>, Helun Song<sup>d</sup> \*

Suzhou Institute of Nano-Tech and Nano-Bionics, Chinese Academy of Sciences, Suzhou, China

<sup>a</sup>fli2012@sinano.ac.cn, <sup>b</sup>xzzhang2016@sinano.ac.cn, <sup>c</sup>yhzhang2006@sinano.ac.cn, <sup>d</sup>hlsong2008@sinano.ac.cn

**Keywords:** infrared thermal image, guided filter, detail enhancement, ORB, Line Segment Detector

**Abstract:** In order to use infrared thermal imaging for simultaneous localization and mapping of mobile robot vision navigation, a point-line feature extraction and matching algorithm after infrared image enhancement is proposed for visual odometry. We first used a guided filter to smooth the input image and separate it into the base layer and the detail layer. Then constraining the gradient of the detail will be used gain mask to enhance it. Finally, the two parts of the image combined with weighted coefficients will be exported into the second guided filter. The output image will use LSD to extract line feature, then using ORB algorithm will get point feature extraction and matching from line feature image. The results show the it can effectively improve the defect of infrared image blurred with similar background temperature and difficult to extract image features for mobile robot visual SLAM.

## 1. Introduction

Monocular visual odometry for its rich visual information and the computation is relatively small, and become an important branch of Visual-Simultaneous localization and mapping(V-SLAM)[1].

The properties of the infrared thermal imaging technology make it not only visually display the object surface temperature field in the form of pictures, but also enable it to show the environment and identify targets highly robustness to changing lighting conditions and other environmental effects. It is a relatively new field to exploit the thermal infrared imaging as a mobile robot sensing modality, such as being used for mobile robot object Identification [2], or SLAM.

However, compared with the visible light camera, the context of infrared images in the video-based SLAM system has some most serious challenges:

- Limited resolution
- High noise
- Low contrast
- Poor texture distribution

The four listed disadvantages of infrared thermal imaging have a great influence to achieve a large number of reliably tracked features between frames in a thermal-infrared video sequence. Therefore, infrared image pretreatment method is meaningful and necessary for robot vision.

Point feature is the most widely used feature in visual SLAM, and it is mature in feature extraction, matching and representation. However, the point feature is more dependent on the environment and does not perform well in scenes such as missing textures. In the artificial structured environment, there are abundant line features, which play a complementary role with the point features in the image. At the same time, the line feature is a higher level feature, and the environment map constructed by line segment feature has more intuitive geometric information. Because of the thermal infrared images characteristic defects, extracting line features is more needed.

The usual points are specialized as corner points or spots in images, and specific detection algorithms are needed to extract the features of these images Historically, researchers have proposed many image features. The SIFT is a classic, has good scale, rotation, perspective and illumination

invariance, but the large amount of calculation and time-consuming, so far, PC CPU was unable to compute SIFT features. Other features are considered to reduce the accuracy and robustness, and to improve the speed of calculation ORB feature is a representative real-time image feature. It improves the problem that the FAST detector does not have directionality, and uses the fast binary descriptor BRIEF to accelerate the feature extraction of the whole image.

At present, there are many visual odometry and visual SLAM based on line segment features too. Lu Y [3] Used point and line segment features to construct a SLAM system based on key reconnaissance optimization, Guoxuan Zhang [4] proposed the pure line features of SL-SLAM (Stereo Line-based SLAM), framework based on the binocular vision SLAM system, including the construction of line feature map of the scene, closed loop detection and so on.

According to using infrared image for VSLAM and comprehensive efficiency and cost, this paper proposed an algorithm which firstly doing image enhancement and denoising and then deal with point and line feature extraction. This pretreatment process can effectively improve the infrared visual navigation robot localization accuracy and robustness.

## 2. The Principle of The Proposed Algorithm

### 2.1 Guided Image Filtering

Guided Image filter [5] smooths an input image  $p$  under the guided image  $I$ . Local linear model assumes that filtered output  $q$  is expressed by a linear function of  $I$  in a window  $\omega_k$  centered at pixel  $k$ :

$$q_i = a_k I_i + b_k, \forall i \in \omega_k \quad (1)$$

where  $(a_k, b_k)$  are some linear coefficients assumed to be constant in window  $\omega_k$ . A square window is used in the original formula of 2D guided filtering. We can determine the linear coefficients by minimizing the difference between  $q$  and  $p$ .

$$E(a_k, b_k) = \sum_{i \in \omega_k} ((a_k I_i + b_k - p_i)^2 + \epsilon a_k^2) \quad (2)$$

The least squares solution of (2) is given as:

$$a_k = \frac{\frac{1}{|\omega|} \sum_{i \in \omega_k} I_i p_i - \mu_k \bar{p}_k}{\sigma_k^2 + \epsilon} \quad (3)$$

$$b_k = \bar{p}_k - a_k \mu_k \quad (4)$$

where  $\mu_k$  and  $\sigma_k^2$  are the mean and variance of  $I$  in  $\omega_k$ ,  $|\omega|$  is the number of pixels in  $\omega_k$ , and  $\bar{p}_k$  is the mean of  $p$  in  $\omega_k$ . Because each pixel lies in the overlapping windows, then a pixel's smooth result is the average of multiple estimates, we rewrite the final Equation:

$$q_i = \bar{a}_1 I_i + \bar{b}_1 \quad (5)$$

where  $\bar{a}_1 = \frac{1}{|\omega|} \sum_{k \in \omega_i} a_k$  and  $\bar{b}_1 = \frac{1}{|\omega|} \sum_{k \in \omega_i} b_k$  are the average coefficients of all windows overlapping  $i$ .

### 2.2 Gain mask enhancement for detail component

The edge-preserving filtering property of guided filter explained intuitively help us to design the gain mask. [5] Consider the case where  $I = p$ .

Case 1: "High variance" .

Case 2: "Flat patch" .

We simplify the gain mask as follows:

$$T = (KH - KL)\bar{a}_1 + KL \quad (6)$$

The parameter  $KH(\text{the max}), KL: (\text{the min})$  can modify the masking function, which makes the detail enhancement more controllable. With this function we are able to enhance the detail region as much as we need, and directly using  $\bar{a}_1$  save a number of computation.

### 2.3 Line Segment Detector

LSD (Line Segment Detector) [6] is an image processing method for fast detection of line segments in images. It was Published by Rafael Grompone von Gioi et al in 2012. It is a local extraction line segment which can obtain sub-pixel precision, and the output of LSD results is the coordinates and attributes of the line segment, including the starting point, the end point, the thickness of the line, and so on.

The first step of LSD is to scale the input image to 80% of its size. Then, the gradient magnitude and gradient direction of the image are calculated. The gradient field LSD is formed by all pixel gradients. Firstly, the angle between each pixel and the level-line is calculated to form a gradient phase field (level-line Field). Merge the pixels with approximately the same direction in the field, so that a series of regions can be obtained, these regions are called line support regions. Fig 1.

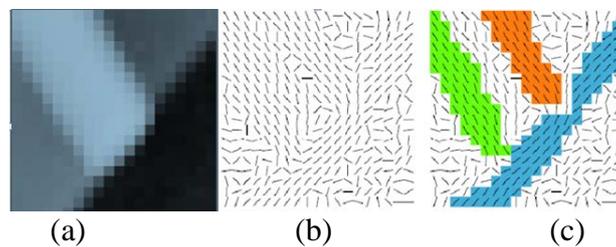


Figure 1. LSD line detection. (a) gray image. (b) level-line Field. (c) line support regions.

Finally, to verify each rectangular area, the main direction of a support line of pixels in the area of the level-line angle and the minimum bounding rectangle of the angle difference in tolerance (tolerance) in  $2\tau$ , then this point is called aligned point". The minimum number of pixels in the circumscribed rectangle and the number of aligned points in it are used to determine whether the line support region is a straight line segment. When the ratio is greater than or equal to the threshold  $D$ , the Number of False Alarms function NFA is calculated. If the NFA is less than or equal to the internal parameters, the line segment is identified as the line segment.

### 2.4 ORB feature extraction

ORB (Oriented FAST and Rotated BRIEF) [7], the algorithm is improved based on the BRIEF feature description algorithm and the FAST corner extraction algorithm.

FAST is defined that if a pixel is located in different regions with enough pixels in its surrounding area, the pixel may be an angular point. Compared with other corner detection algorithms, FAST only needs to compare the pixel brightness size. When a discrete Bresenham circle with a radius equal to 3 pixels is taken as the center of the pixel, there are 16 pixels on the boundary of the circle, Fig. 2.

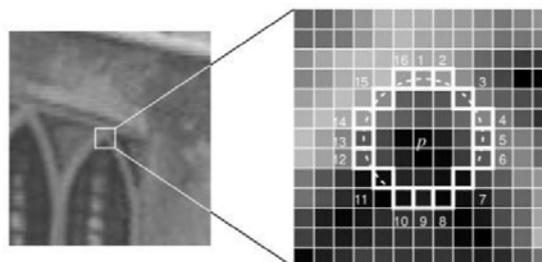


Figure 2. FAST Characteristic point

To select a pixel  $P$  from the image, we first set the brightness value to  $I_p$ . Set a proper threshold  $T$  (for example,  $I_p$ 's 20%). If there are  $n$  continuous pixels in the size of 16 pixels on a circle, or their pixel values are larger than those of  $I_{p+t}$ , or smaller then  $I_{p-t}$ , then it is a corner.

In view of the fact that the FAST corner has no directionality and scale weakness, the gray centroid method is proposed. Connecting the geometric center O and centroid C of the image block, then the angle between the feature point and the centroid is defined as the direction of the FAST feature point. FAST corner has the description of scale and rotation, which greatly improves the robustness of the expression between images. So in ORB, the improved FAST becomes Oriented FAST.

BRIEF, which provides a binary descriptor which consists of many 0 and 1, does not need to compute a feature descriptor similar to SIFT. It needs to smooth the image first, then select a Patch around the feature point, and pick a pair of n points in the Patch by a selected method. Then, for each pair of points (p, q), we compare the brightness values of these two points, and if  $I(p) > I(q)$ , then this point pair. BRIEF uses a randomly selected, very fast, and the use of the binary expression, storage is also very convenient, generally speaking, BRIEF is suitable for feature extraction of a very high efficiency in real-time image matching method.

### 3. Experimental

#### 3.1 The process of two Guided filter enhancement method

The workflow of the method displays in Fig.3. In Fig. 3(a) is the original IR image. Fig. 3(b) and (d) are the base layer and detail layer after filtering original image by GF. Then, we first project histogram the base layer's gray levels distribution and get Fig. 3(c), and it can be noticed that the brightness of Fig. 3(c) is brighter than Fig. 3(b). Meanwhile, we process the detail layer. As shown in the detail layer Fig. 3(d), there is very dark and blurry of the details because the temperature of objects in original IR image is very small. It is necessary to make sure that the detail derivatives and the input derivatives have the same sign. After applying the technique removal of halos artifacts, the correction result in Fig. 3(f). Then the gain mask enhancement of Fig. 3(f) is shown in Fig. 3(g). After dealing with the base layer and detail layer, Fig. 3(e) is the combination of them. Finally, Fig. 3(h) is the output of the second GF of Fig. 3(e).

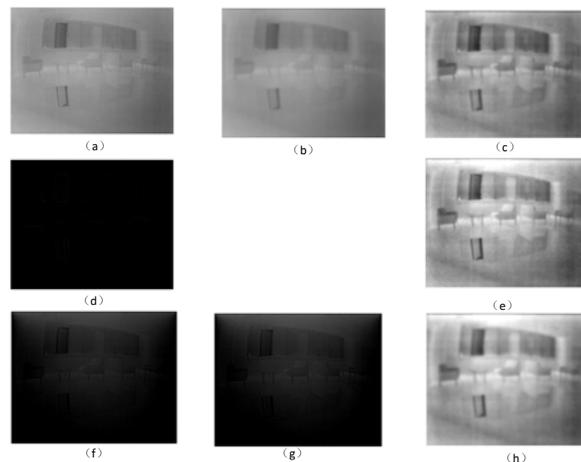


Figure 3. The proposed method flow chart on test image . (a) Original IR image. (b) The base layer by GF. (d) The detail layer by GF. (c) The projection histogram of base layer. (f) The halo removal of detail layer. (g) The Gain mask enhancement of (f). (e) The combination of (c) and (g). (h) The second GF of (e).

The bright of the enhancement could see the details clearly. The contrast is better than original images, and the local details are perceived well. The parameters setting are as follow: the twice of GF is same as:  $r = 4$ ;  $\text{eps} = 0.5^2$ . The Gain Mask is  $KL = 1$ ;  $KH = 3$ ;

#### 3.2 Feature point extraction and matching of ORB

As shown in figure 4~5, the experiment shows that the ORB feature point matching has good

stability in the image with obvious heat source, and is not weaker than that of the visible image with rich feature points. However, if the intensity of image enhancement is not enough, the possible feature points are scarce in the background environment with similar temperature. It is easy to lose when running ORB-SLAM, aiming at this problem. A strengthening of image enhancement is not the same, there will be a lot of noise also increases, there will be a large number of false feature points, and also easy to frame loss. So the stability of feature information needs to be improved from other aspects, such as the development of new feature lines or semantic maps.

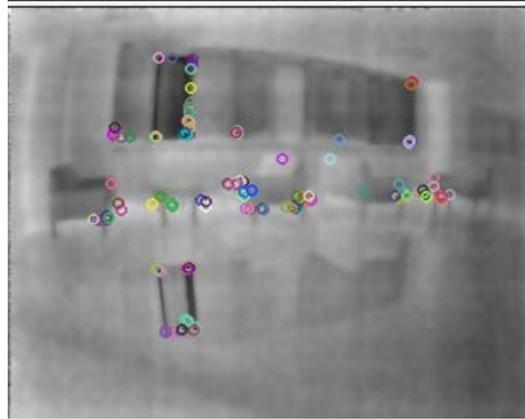


Figure 4. Extracting ORB feature points from infrared images with detail enhancement

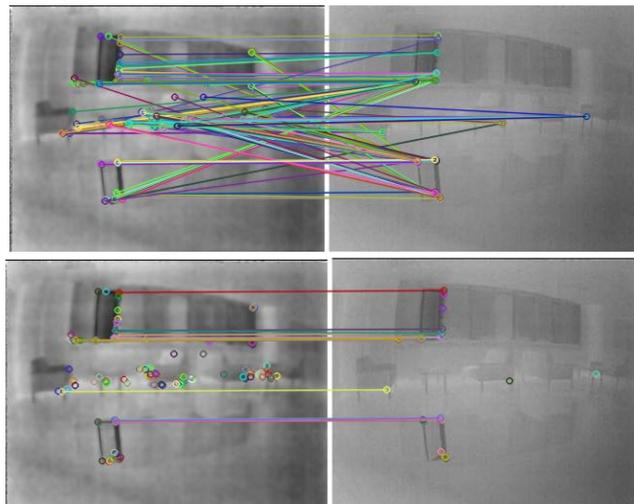


Figure 5. feature points matching selection from diferent enhancement

### 3.3 LSD line extraction

LSD algorithm is used to extract the line features of infrared image.

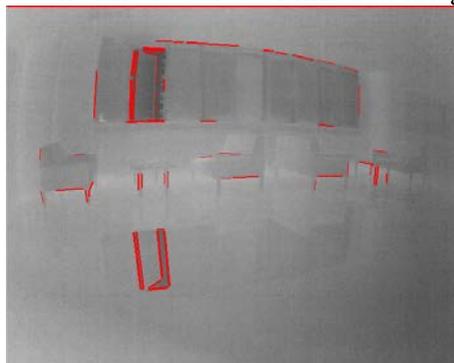


Figure 6. using LSD to extract line features

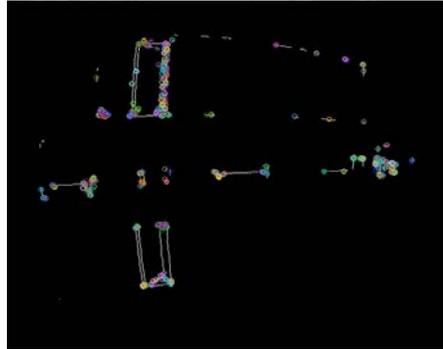


Figure 7. Extracting ORB feature points from line feature image

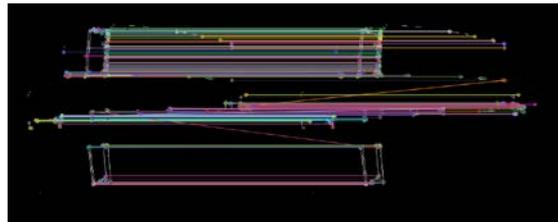


Figure 8. feature points matching selection in line feature image

In our work, after the extraction of line features in the actual process, it can extract the matching points with ORB algorithm, convergence to the ORB-SLAM system, aiming at the basic background temperature as cause extremely lack of scene texture.

#### 4. Conclusions

This paper studies how to develop and use the infrared thermal imaging for mobile robot localization and mapping as visual vision sensor, dig the characteristics of infrared image, infrared image with respect to the defects of the visible image denoising, feature extraction, enhancement, increasing the line features aspects of strengthening the robustness and meet the real-time constraint down achieve the aim of study through the images, line features can make up for the lack of feature point feature scenes, especially for the infrared image background temperature often very similar and very lack of texture, to further enhance the robustness of the front end of SLAM. Improvement the effect of each module is verified by experimental test.

#### References

- [1] G. Klein and D. Murray. "Parallel tracking and mapping for small AR workspaces," In International Symposium on Mixed and Augmented Reality (ISMAR), 2007.
- [2] Fehلمان Ii, William L., and M. K. Hinders. "Passive infrared thermographic imaging for mobile robot object identification," *Journal of Field Robotics* 1211.1(2010):2076-2083.
- [3] Lu Y, Song D. Robust RGB-D Odometry Using Point and Line Features[C]// IEEE International Conference on Computer Vision. IEEE, 2015:3934-3942.
- [4] Zhang G, Lee J H, Lim J, et al. Building a 3-D Line-Based Map Using a Stereo SLAM[J]. *IEEE Transactions on Robotics*, 2015:1-14.
- [5] K. He, J. Sun, X. Tang, "Guided image filter," *IEEE Trans. Pattern Anal. Mach. Intell.*35 (6) (2013) 1397–1409.
- [6] Gioi R G V, Jakubowicz J, Morel J M, et al. LSD: A line segment detector[J]. *Image Processing on Line*, 2012, 2(4):35-55.
- [7] Rublee E, Rabaud V, Konolige K, et al. ORB: An efficient alternative to SIFT or SURF[C]// IEEE International Conference on Computer Vision. IEEE, 2012:2564-2571.