

Weighted Multi-Scale Image Matching Based on Harris-Sift Descriptor

Can Sun, Jin-ge Wang, Zaixin Liu, Junmin Li*

School of Mechanical Engineering and Automation
Xihua University, Chengdu, China

*Corresponding author, e-mail: lijunmin1975@163.com

Abstract

According to the rotational invariance of Harris corner detector and the robustness of Sift descriptor. An improved Harris-Sift corner descriptor was proposed. At first, the algorithm given multi-scale strategy to Harris corner, improved corner counting method and removed redundant points at the same time, then, the corner was directly applied to low-pass Gaussian filter image. Based on the histogram of Sift feature descriptor, generates a new 128-dimensional feature vector descriptor by multi-scale Gauss weighted. Through the above, Harris corner detector and Sift descriptor was normalized in the scale layer and gradient features. The experiment results indicated that, the improved corner descriptor comprised both advantage of Harris corner detection and Sift feature descriptor. The method reduced the computation time and the false match rate, which could be validly applied to the robot stereo vision matching and three-dimensional reconstruction.

Keywords: stereo vision, robot, corner detection, feature descriptor, scale

Copyright © 2013 Universitas Ahmad Dahlan. All rights reserved.

1. Introduction

Computer stereo vision technique mainly includes five parts [1]: image acquisition, feature extraction, camera calibration, stereo matching and 3D reconstruction. In these modules, the correct match of different image feature is necessary for robot cognize space objects, this process is called feature points matching [2]. The impact of the light in the pairing of the feature points, the correct description of the feature points in different scale space and image rotation changes base coordinate system are all as the inevitable role in image matching process.

For the influence of the above factors on the matching, feature points described in the present method comprises: Differential-based descriptor [3], Descriptor based on moment invariants [4], Distribution-based descriptor. Koenderink made a deep research on the image local differential properties, and puts forward the concept of scale parameter [5]. Freeman specified a controllable direction differential filter [6], the filter effectively avoided the impact of image rotation feature point matching. Mikolajczyk. K made a comprehensive summary of the descriptor algorithm [7] who found that Lowe's Sift operator [8] with a strong stability, the algorithm built DOG scale space at first, got image essential features characteristics. Then calculated the extremum point by using Hessian matrix, given gradient attribute to extremum point made it become vector feature points, finally, precise classified histogram information of each feature point, it ensuring the uniqueness of feature point matching.

It is inconvenient to calculate the Hessian matrix in Sift point detection, 128-dimensional corner feature vector descriptor need a long time to judge. It resulting in the Sift algorithm does not have the real-time of determine corner and endows the ability to the descriptor. This paper combined the Harris corner detector to the Sift algorithm, first, expanded scale space on Harris corner detection, completed consistency of Harris corner detection and Sift in the scale parameter, because the Harris corner has rotation invariance, it can be completed replace the Sift extreme points in the detection process, by using the Sift operator give the gradient feature to the Harris corner directly, built a stable feature points which does not vary with image rotation, scaling, affine transformation, illumination changes and other factors. And generated point matching fused image with image fusion algorithm.

2. Detect Interest Point

In order to extract the required points, selected Harris corner detector based on gray level changes. Detect average energy changes in the larger image area while the detector is running. If it changes at x, y direction, considered to be the corner, if only in one direction changes greatly, considered to be the edge feature, if the changes in both direction is not obvious, for imagesmoothing region.

2.1. Improved Harris Corner Scale Space

Harris corner is determined only by the practical experience of the threshold to judge whether metric function was a corner or not. Corner autocorrelation matrix [9] as follows:

$$\mathbf{M} = \mu(\rho, \sigma_N, \sigma_D) = \sigma_D^2 \mathbf{g}(\rho, \sigma_N) \otimes \begin{bmatrix} L_x^2(\rho, \sigma_D) & L_x L_y(\rho, \sigma_D) \\ L_x L_y(\rho, \sigma_D) & L_y^2(\rho, \sigma_D) \end{bmatrix} \quad (1)$$

where ρ is a Gaussian window position, $L_x(\rho, \sigma_D)$, $L_y(\rho, \sigma_D)$ respectively are horizontal and vertical differential operator at a position ρ , $L_x L_y(\rho, \sigma_D)$ is per-pixel operations. Type (1) had already replace the original Harris Gauss window into discrete 2D zero-mean Gauss function $\mathbf{g}(\rho, \sigma_N) = \exp\left(-\frac{(x^2 + y^2)}{2\sigma^2}\right)$, defining $\sigma_N = \xi^n \sigma_0$, σ_0 is the initial scale value, $\sigma_0 = 1.0$,

ξ^n is the ratio of the adjacent scales, normally $\xi = 1.4$, n for scale layers, generally take $n = 7$, \mathbf{M} as the autocorrelation matrix of target point within the Gaussian window.

By replacing multi-scale Gaussian filter $\mathbf{g}(\rho, \sigma_N)$, the Harris detector with single scale become multi-scale spatial detector with linear growth which in comparison of the scale space aspect.

2.2. Defined Harris Corner Response Function

It can get the corner response function [10] by calculating the characteristic values λ_1 , λ_2 of the correlation matrix \mathbf{M} :

$$\mathbf{R} = \det(\mathbf{M}) - k * \text{trace}^2(\mathbf{M}) \quad (2)$$

coefficient $k = [0.04, 0.06]$, the value directly impact on the result of function response. In order to stabilize corner amount in each time series sample pictures, improved corner response algorithm as follow:

$$\mathbf{R} = \frac{L_x^2 * L_y^2 - (L_x * L_y)^2}{L_x^2 + L_y^2} - k * \text{trace}^2(\mathbf{M}) = \text{cim} - k * \text{trace}^2(\mathbf{M}) \quad (3)$$

$R > 0$, expressed as a corner feature, when $R < 0$, expressed as edge feature, when R is very small, smooth area that is within the window. $\text{trace}(\mathbf{M})$ is the trace of \mathbf{M} . $\text{trace}(\mathbf{M}) = \lambda_1 + \lambda_2$ as corner detection process data which describe the local features of an image. Figure 1 (a) is a test image, Figure 1 (b) to the corresponding trace image.

2.3. Remove Redundant Points

Figure 1 (b) produced duplicate traces when the scale space increase, the existence of repeat traces reacted duplicate detection of corner, optimization eliminate redundant features point as follows:

$$\sigma(x_i) = \left\{ \max(\sigma(x_i), \sigma(x_k)) \mid \|x_k - x_i\|^2 \leq 2 \right\} \quad (4)$$

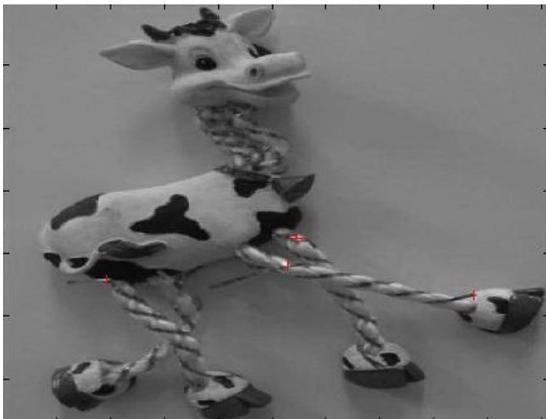


Figure 1(a). Test image

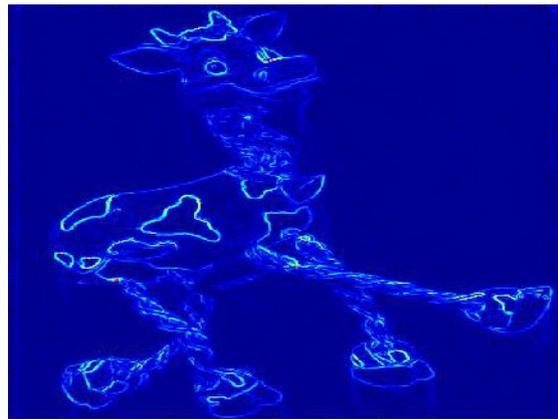


Figure 1(b). Trace image

Equation (4) can avoid image minutiae detection failure in large-scale effectively and detect large range of the image feature point in the small scale, which x_i , x_k correspond to the low scale and high scale Gaussian space.

The above improvements Harris algorithm given the characteristics of multi-scale to detector, it provides accurate corner points to image feature extraction and image matching.

3. Improved Harris-Sift Corner Descriptor

Sift algorithm mainly includes DOG layered scale space construction, image pixel extreme point detection, determine the scale invariant feature points, remove the unstable feature points, determine the interest principal direction of reliable points, generating feature descriptor. Apply this feature matching algorithm can effectively overcome imaging plane translation, scaling, rotation and affine transformation in the images matching

3.1. Create Gradient Characteristics

The Sift descriptor used non-maxima suppression algorithm [11] compared all extreme points in the same scale layer and the upper-lower scale layer, resulting in a large amount of calculation. In order to avoid the problem of real-time lack in Sift algorithm, we take multi-scale Harris corner instead of Sift extreme points, using sift key point allocation method, calculate gradient magnitude and direction of new Harris point:

$$\begin{cases} m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \\ \theta(x, y) = \tan^{-1} \left(\frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)} \right) \end{cases} \quad (5)$$

(5) is the corner gradient attribute. Notice that $L(x, y)$ for each low-pass Gaussian filter image of the key point, the improved Harris corner will be key point in the calculation, while the scale of improved Harris point as the key point scale layer. Then sample in the feature points within the neighborhood window and creates a gradient direction histogram of the feature points in this window, we selected improve Harris corner Gaussian window $g(\rho, \sigma_N)$, the window size wastaken to be 1.5 times σ_N . In the histogram which exists 80% remaining of main peak gray column, it can define the direction of this gradient to the auxiliary direction of the feature point. Each feature point with only one main direction, but there is more than one auxiliary direction. Finally, make each key point get three characteristics: scale, direction and position.

3.2. Generate New Harris-Sift Feature Descriptor

Ensure the rotation invariance of the descriptor at first. Select 2×2 Gaussian window on the point, gradient magnitude and direction in each window was calculated by using formula (5). Sift algorithm recommended 4×4 Gaussian window on the point, we can introduce a Gauss distribution function, weighted distance feature points, then, each of the key points include 128 dimension, normalized the 128-dimensional vector, which generate new Harris-Sift feature point descriptor.

3.3. The Matching Criterion of Image Feature Points

Generally the coincidence rate of different pictures in same scene is more than 50%, using simple threshold-based matching:

$$\tau = \sqrt{m_p^2 + m_q^2} \quad (6)$$

m_p , m_q respectively are the descriptor distance for two feature points in the same scene objects which corresponding to the set of $P\{p_1, p_2, p_3 \dots p_n\}$, $Q\{q_1, q_2, q_3 \dots q_n\}$, τ is descriptor two-dimensional Euclidean distance, $\tau = [0.4, 0.6]$.

Thus, feature point descriptor can be matched with different sets of points at the same scale invariant and rotation invariant, and to distinguish false matching points effectively. The algorithm block diagram of binocular visual feature point descriptor is shown in Figure 2.

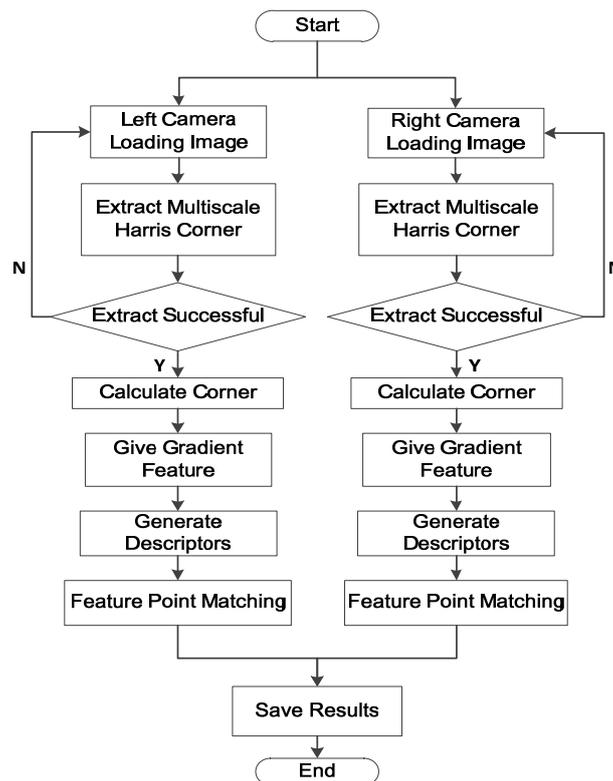


Figure 2. Feature matching algorithm diagram

4. Experimentation

Using MATLAB programming environment to working the algorithm in Core E5500 2.8GHZ dual-core CPU, 2GB RAM computer. Figure 3 is the binocular camera as the image capture device. Each piece of experimental images pixel ratio is 640×480 , bit depth is 8bit.



Figure 3. Binocular vision system

4.1. Camera Calibration

Calibrate the camera at first, the parameters obtained by plane calibration target [12] calculation:

$$\mathbf{F} = \begin{bmatrix} f_x & 0 & u_0 \\ 0 & f_y & v_0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 831.88690 & 0 & 321.20991 \\ 0 & 830.19993 & 241.56118 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{R} = [r_1 \quad r_2 \quad r_3] = [1.95 \quad 1.78 \quad -0.15]$$

$$\mathbf{T} = [t_x \quad t_y \quad t_z] = [-23.24 \quad -27.42 \quad 225.31]$$

$$\mathbf{K} = [k_1 \quad k_2] = [-0.42 \quad 0.19]$$

$$\mathbf{P} = [p_1 \quad p_2] = [-0.0003 \quad -0.0005]$$

$$\boldsymbol{\varepsilon} = [\varepsilon_1 \quad \varepsilon_2] = [0.0004 \quad 0.0003]$$

\mathbf{F} for the matrix of camera intrinsic parameters, f_x , f_y respectively is component of a lens scale focal length about x-axis and y-axis in the image coordinate system, u_0 , v_0 for the center coordinate of the image. \mathbf{R} , \mathbf{T} are the external camera parameter matrix. \mathbf{K} , \mathbf{P} , $\boldsymbol{\varepsilon}$ for three distortion coefficients of camera intrinsic.

4.2. Different Attributes of Single Image Matching

Compared with the SIFT algorithm, matching two pictures feature point which different in the brightness unevenness and pixel ratio respectively, taking $\tau = 0.6$ in (6) to ensure that is able to detect a stable feature point.

Test matching effect of two images for luminance unevenness in the same scene, the processing result in Figure 4 (a), Figure 4 (b) and Table 1 compared two methods of match time, the number of corner detection, the number of feature points matching and the matching rate.

Test matching effect of two images for different pixel ratio in the same scene using the same way, the processing result in Figure 5 (a), Figure 5 (b) and table 2 compared two methods.

The comparative data in Table 1 and Table 2 shows matching stability of improved feature points, reduced the complexity of computing time, improved the detection precision of the corner points and the number of correctly paired, descriptor has well stability and real-time.

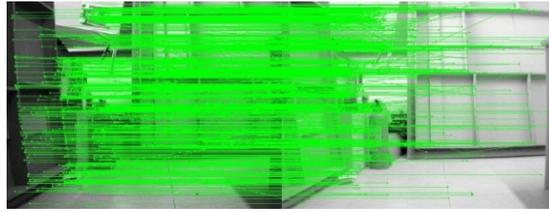


Figure 4 (a). Sift descriptor in illumination changes match



Figure 4 (b). Sift descriptor in illumination changes match

Table 1. Algorithm comparison

Descriptor	Time (s)	Corner amount	Pairing number	Matching rate
Sift	9.8	1825	586	90.0%
Improve	4.2	1136	347	90.5%

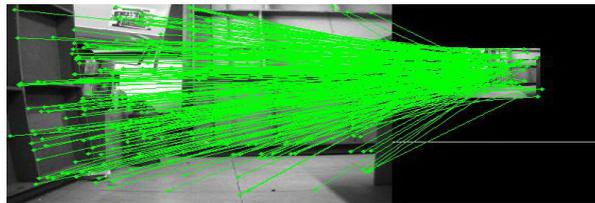


Figure 5(a). Sift descriptor in different pixel ratio

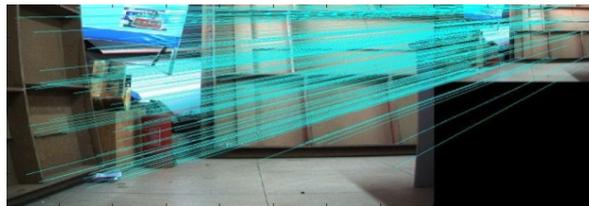


Figure 5(b). Improve descriptor in different pixel ratio

Table 2. Algorithm comparison

Descriptor	Time (s)	Corner amount	Pairing number	Matching rate
Sift	12.5	580	423	96.4%
Improve	7.2	317	240	96.9%

4.3. Binocular Visual Matching and Fusion

Then, the algorithm is applied to binocular stereoscopic feature point matching, the results shown in Figure 6 (a) and Figure 6 (b). In the image matching, image fusion algorithm

[13] was introduced, defined fusion factor $a = \frac{(x, y) - (x_{\min}, y_{\min})}{(x_{\max}, y_{\max}) - (x_{\min}, y_{\min})}$, Computed the change

process of boundary coordinates x_{\min}, y_{\min} to x_{\max}, y_{\max} , matched up the repeating feature C between image A and image B, $C=A \cap B$. $P(i, j) = a(j) \times A(i, j) + (1 - a(j)) \times B(i, j)$ is the mixing pointgradation of synthetic image D, Figure 7 is the binocular visual features of the image fusion results.



Figure 6 (a). Sift descriptor with binocular visual



Figure 6 (b). Improve descriptor with binocular visual



Figure 7. Image fusion

It can be seen that the matching featureline of two images by almost horizontal without epipolar line constraints. It has good results in the center region fusion about two pictures. The algorithm can be used in the binocular stereo matching.

5. Conclusion

The method takes the advantage of Harris corner detection in quick and simple and Sift feature descriptor in accurate and stable. Considering the impact of scale space and Gauss weighted on computational efficiency, eliminate redundant corner, reasonable evaluate feature point matching criterion, calibrate the parameters of camera so as to correct image pincushion distortion, match up the repeating feature with image fusion algorithm. The experiment results shows that the improved Harris-Sift feature point descriptor could effectively avoid the condition of false matching within illumination, scaling, translation, affine transformation which consist in

the matching procedure. It is an efficient and stable algorithm can be used in robot vision system.

Acknowledgements

We would like to appreciate the supported by the Educational Commission of Sichuan Province of China (12ZB129), the Open Research Fund of Key Laboratory of Manufacture and Automation Laboratory (Xihua University, Szjj2012-022), Innovation Fund of Postgraduate (ycjj201276) of Xihua University for the first author.

References

- [1] Budiharto Widodo, Santoso Ari, Purwanto Djoko, Jazidie Achmad. Multiple moving obstacles avoidance of service robot using stereo vision. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2011; 9(3): 433-444.
- [2] Radke RJ, Andra S, Al-Kofahio, et. al. Image change detection algorithms A systematic survey. *IEEE Transaction on Image Processing*. 2005; 14(3): 294-307.
- [3] Yu Dao-li. Image Matching Algorithm Based on the Local Invariant Features. Master Thesis. Anhui: Department of computer science and engineering of Anhui University; 2012.
- [4] M David. Vision. San Francisco: Freeman Publishers. 1982.
- [5] JJ Koenderink, AJ van Doorn. Representation of local geometry in the visual system. *Biological Cybernetics*. 1987; 55: 367-375.
- [6] WT Freeman, EH Adelson. The design and use of steerable filters. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1991; 13: 891-906.
- [7] K Mikolajczyk, T Tuytelaars, C Schmid, et. al. A comparison of affine region detectors. *International Journal of Computer Vision*. 2005; 65(1/2): 43-72.
- [8] DG Lowe. Distinctive image features from scale-invariant keypoints. *International Journal on Computer Vision*. 2004; 60(2): 91-110.
- [9] C Harris, M Stephens. *A combined corner and edge detector*. Proceedings of the 4th Alvey Vision Conference. 1988: 147-151.
- [10] Zhen-Ying Xu, Wen-Bin Li, Yun Wang, Chun Luo, Shu-Yuan Gao, Dan-Dan Cao. Trinocular calibration method based on binocular calibration. *TELKOMNIKA Indonesian Journal of Electrical Engineering*. 2012; 10(6): 1439-1444.1
- [11] Wang Jing, Wang Hai-liang, Xiang Mao-sheng, Wei Li-deng, Liu Zhong-sheng. Subpixel accuracy central location of circle target based on nonmaximum suppression. *Chinese Journal of Scientific Instrument*. 2012; 33(7): 1460-1468.
- [12] Zhao Xue-feng, Zhong Xiao-min, Lan Yi-hua. Improvement of Tsai's camera calibration under coplanar calibration board nearly parallels to imaging plane. *Computer Engineering and Design*. 2011; 32(3): 1019-1021.
- [13] Tao Zhi-jiang, Huang Hua. Fundus images mosaic based on improved harris and SIFT algorithm. *Computer Engineering and Design*. 2012; 33(9): 3510-3511.