

Study on Improving Image Feature Points Detection and Matching Accuracy in Binocular Vision System

Gang Tong^a, Changcheng Wang^b, Pan Wang^c

Department of Automation and Electrical Engineering, Qingdao University of
Science & Technology, Qingdao, 266042, China

^aemail:tg.gg@163.com, ^bemail:530940299@qq.com, ^cemail:1017599876 @qq.com

Keywords: Image Matching; Feature Points Detection; SIFT Algorithm; Binocular Vision

Abstract. Image feature points detection and matching is the key to binocular vision system performance. The paper is to improve its matching accuracy. In the experiments, Harris algorithm, Susan algorithm and CSS algorithm were used on the same image to extract feature points. Compared with each other, three methods showed different advantages in terms of extracting feature points. And two methods were carried out in the feature points matching process, one method was based on Harris feature points detection while another method was based on SIFT algorithm. The results showed that SIFT algorithm had better matching effect, but matching accuracy remained to be further improved. As a result, we extended the search scope of the extreme points in DoG scale space of the SIFT algorithm and removed feature points around image boundary. Though the number of the detected points changed little, but its detecting accuracy was more reliable. Compared with the effect of traditional SIFT algorithm, the matching accuracy has been significantly improved.

Introduction

Binocular vision is an important part in machine vision system and its idea is to use two same CCD cameras taking photos in a fixed position. Get the depth information of the scene by computing the parallax of the same feature points in two images. With the development of image processing technology, dimensional measurement, 3D reconstruction and target tracking and identification based on binocular vision have become hot research areas.

Image feature points detection and matching are important parts in binocular vision system. There are some good feature point detection methods, such as Harris algorithm [1], Susan Algorithm [2], CSS algorithm [3] and etc. Harris algorithm has good corner extraction capability and strong anti-rotation capability, but it doesn't resist scaling and noise. Susan algorithm is good at resisting noise. CSS algorithm has a good location in detail scale, but the noise also increases accordingly. SSD algorithm, NCC algorithm, RANSAC algorithm and SIFT algorithm are good image matching algorithms [4]. Their merits depend on the matching accuracy and the ability to resist influences. Improvement on SIFT algorithm becomes a hot research in recent years.

In the experiment, Harris algorithm, Susan algorithm and CSS algorithm were carried out. Simulation showed that Harris relatively had the best result. But image matching based on Harris and NCC algorithm was not better than SIFT algorithm. And we extended the search scope of the extreme points in DoG scale space of the SIFT algorithm. Matching accuracy could achieve better results.

Feature Points Detection

Feature points are important image information. They are located in somewhere existing gray level jump and curvature upheaval. And feature points detection is the precondition of image matching. Common feature point extraction methods have Harris, Susan and CSS feature detection algorithms, all these algorithms have their own advantages.

Harris algorithm is to select a target pixel as the center of the window in the neighborhood of one

feature point. Calculate the value of gray scale changes in four different directions and choose the minimum as the pixel's response function to compare with threshold. If the minimum is bigger than threshold, the pixel is the feature point. Harris algorithm is to analyze gray level changes in any direction.

Susan algorithm is based on the theory that pixels associated with each image point of a local area having the same image brightness. The operator uses a circular template to detect feature points and circular template is isotropic.

CSS algorithm extracts the image edge contour curve, fills curve discontinuity to get a continuous edge curve, calculates the curvature of each pixel in the continuous curve and chooses the point with the maximum curvature as a feature point.

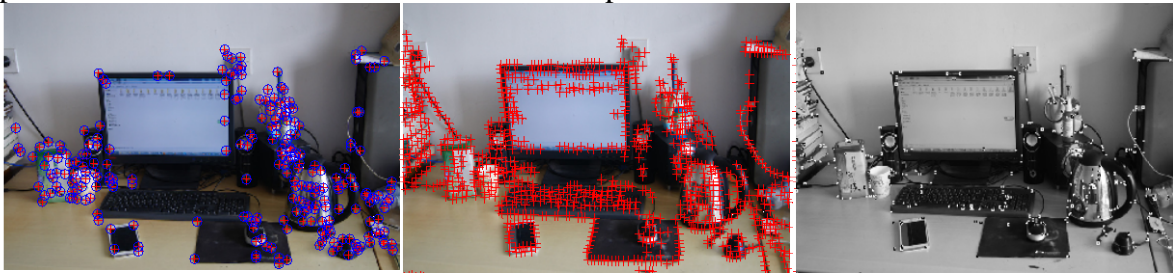


Fig.1. Feature points detection based on Harris, Susan and CSS algorithm

All in all, Harris algorithm has better corner extraction capability and strong anti-rotation capability, but it doesn't resist scaling and noise. Susan algorithm is good at resisting noise. CSS algorithm has a good location in detail scale, but the noise also increases accordingly.

Image Matching

In order to improve the efficiency of feature points matching and find an image matching method having considerable ability in detecting feature points (resisting scaling, rotation and noise), David G. Lowe[5] summed previous experiences and proposed an image matching algorithm based on scale-invariant feature transform (SIFT) [6].

This algorithm calculates the Euclidean distance of points in two images in final matching process. It's a traditional method and the key is to generate the feature vector of each feature point. So researches on its optimized algorithm become hot in recent years.

Specific algorithm process is as follows:

1. DoG scale space: The function, $L(x, y, \sigma)$, shows the scale space of the experimental image and a convolution produced by a variable-scale Gaussian function, $G(x, y, \sigma)$, and an input image, $I(x, y)$:

$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (1)$$

Where $*$ is convolution operation, and

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2} \quad (2)$$

On detection of the feature point in scale space, we could compute the differences of two nearby scales to get the response function, $D(x, y, \sigma)$. Locate the feature points by building Gaussian pyramid and using non-maximum suppression on $D(x, y, \sigma)$.

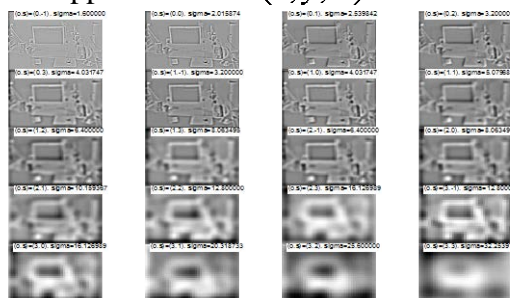


Fig.2. The images of different resolution zoom to the same size

2. Feature points detection and accurate location: Local extreme point is the feature point in DoG scale space. The feature point is not only the extreme point in 3×3 image space, but also is the extreme point in 3×3 scale space. That is to say, the point has information both in image position and scale space.

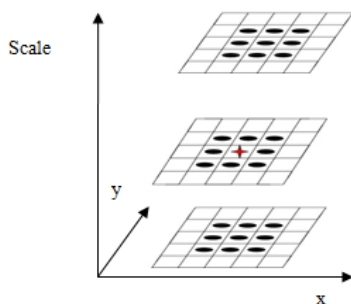


Fig.3. The extreme point location in DoG

Because DoG operator is sensitive on edge response, so further processing, removing points with edge response and low contrast, on extreme points is needed in order to get stable SIFT descriptor.

3. Orientation of feature point: Calculate the image gradient to assign a primary direction for the feature point in order to achieve invariance. Thus, three pieces of information, (x, y, σ, θ) , with each feature point are obtained: the location, scale and orientation.

4. Key point descriptor: According to Lowe's method, the 16×16 neighborhood around the feature point should be divided into 16 blocks and each block is 4×4 . Calculate the gradient in eight directions to get $16 \times 8 = 128$ dimensional feature vector.

5. Image matching: Calculate Euclidean distance of the 128-dimensional feature vector of all key points in two images. Find the closest neighbor and the second-closest neighbor. Two key points are matched successfully when the ratio is greater than or equal to a threshold value (obtained by experiments).

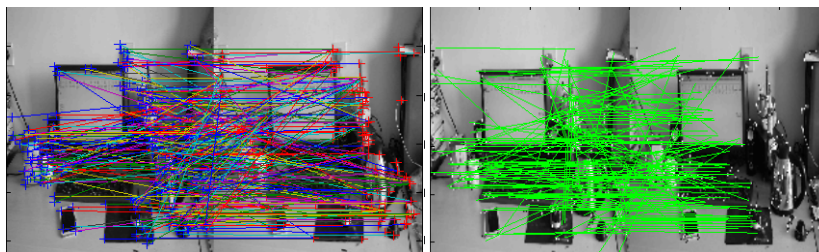


Fig.4. Image matching effects based on Harris algorithm and SIFT

Extend Search Scope

In order to provide more valuable references for feature points location, we extended the search scope from 3×3 neighborhood to 5×5 neighborhood in the experiment. Under the scale space, considered the impact of calculation amount and some uncertainties, we didn't extend the scale space. As a result, the point could compare with 24 points in the same scale and 50 in two nearby scales to ensure that extreme points could be detected both in the scale space and two-dimensional image space.

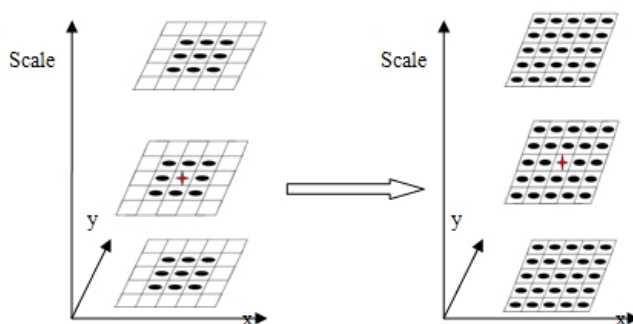


Fig.5. Extend search scope from 3×3 neighborhood to 5×5 neighborhood

Extreme point detection and location process in scale space is as follows:

1. Get the scale space $L(x, y, \sigma)$ by a convolution operation between Gaussian function, $G(x, y, \sigma)$ and an input image, $I(x, y)$. And represent it by a Gaussian pyramid.
2. The Gaussian pyramid subtracts the nearby one to generate differential Gaussian pyramid, $D(x, y, \sigma)$.
3. Detect extreme points in 5×5 neighborhood in the same scale and two nearby scales in differential Gaussian pyramid.
4. Remove points with edge response and low contrast. In order to get better matching effect, we also removed points around the image boundary.

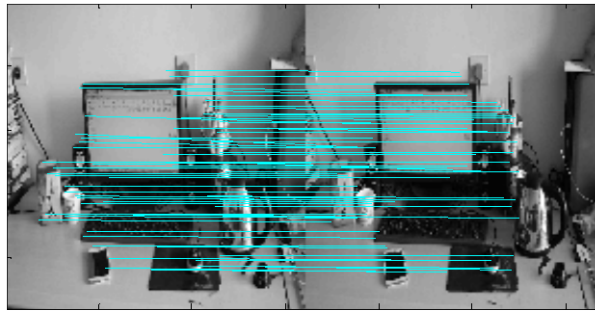


Fig.6. The effect of the optimized SIFT matching

Conclusion

Scale-invariant feature transform (SIFT) is a descriptor for image processing. This descriptor has scale invariance. It is a local feature descriptor and it could detect feature points accurately. Extending the search scope from 3×3 neighborhood to 5×5 neighborhood could get better matching effects than traditional SIFT algorithm and some other matching methods, such as, NCC matching algorithm. The experimental results show that matching accuracy based on Harris algorithm is 69.3%, SIFT method is 89.4% and the optimized method could achieve 97.5% matching accuracy. Improving the positioning accuracy of the feature points could effectively improve the matching accuracy. And the performance of binocular vision system could be also further optimized.

References

- [1] CHENG Bang-sheng, TANG Xiao-we. Study of Harris scale invariant keypoint detector [J], Journal of Zhejiang University (Engineering Science), 2009 43(5) 855-859.
- [2] Smith S M, Brady J M. SUSAN-a new approach to low level image processing. International Journal of Computer Vision, 1997 23(1) 45-78.
- [3] Mikolajczyk K., Schmid C.. An affine invariant interest point detector [C]. European Conference on Computer Vision (ECCV), Copenhagen, Denmark, 2002. 128-142.
- [4] Schmid C, Mohr R, Bauckhage C. Evaluation of interest point detectors. International Journal of Computer Vision, 2000 37(2) 151-172.
- [5] Lowe, D.G. Object recognition from local scale-invariant features [C]. International Conference on Computer Vision, Corfu, Greece, 1999. 1150-1157.
- [6] Lowe D G. Distinctive Image Features from Scale Invariant Keypoints. International Journal of Computer Vision, 2004 2(60) 91-110.