Study of Micro-vision Calibration Technique Based on SIFT Feature Matching

Tao Hu, Hui-Lan Wu, Guodong Liu

Study of micro-vision calibration technique based on SIFT feature matching

Tao Hu[1], Hui-lan Wu[2], Guodong Liu[1]
[1]Harbin Institute of Technology, Department of Automatic Measurement and Control, Harbin, China
{hutao,lgd}@hit.edu.cn
[2]Research Institute of Shanghai Academy of Spaceflight Technology, No.803, Shanghai, China
whl3003@163.com

Abstract. In the micro-vision system, precise calibration of the pixel equivalent is a prerequisite for accurate visual inspection. Traditional calibration process not only needs to take standard parts as the base, but also has strict requirements on their shape, manufacturing precision, placement and so on. Because of such shortcomings of traditional method, this paper gives an improved SIFT calibration algorithm for the pixel equivalent. Firstly, get the image's characteristic points before and after its' moving, using SIFT algorithm. Then filter the mismatched points through the second filtering algorithm so that the matching accuracy can be greatly improved. Compare moving distances of the characteristic point pixels to the reference distance of optical grating. Then accurate pixel equivalents can be obtained. Experiments show that the calibration algorithm is more accurate than traditional methods. So this algorithm can completely replace the standard-parts method in micro-vision system.

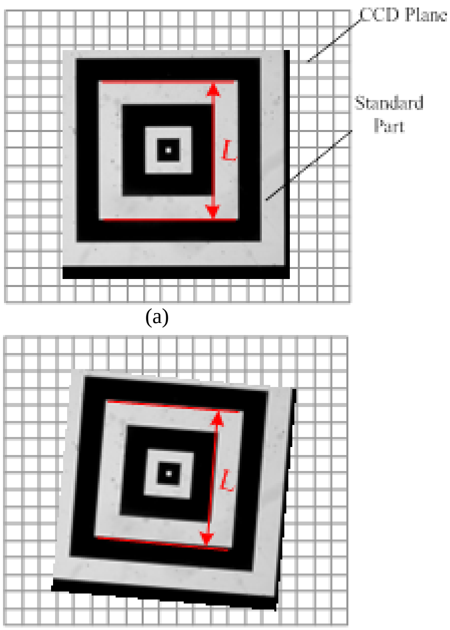
Keywords: Vision measurement; Calibration; Feature Matching

1 Introduction

Measuring object size using CCD pixels as the scale is the core of vision measuring technology. So we need to acquire the actual size of CCD pixel. And because of the existence of optics system, the pixels' actual application size should be the one which is calculated through the optics system. During the application process of the vision system, an accurate enlargement of the optics system is hard to obtain because of many factors, just like the camera's relative position with CCD, and the camera's assembly quality differences. So, calibration of CCD pixel equivalent is the precondition and foundation of the measurement based on image [1~2].

Currently, standard-parts method is commonest among the calibration of CCD pixel equivalent. E.g.: figure 1(a) Standard mask. Test the standard part's geometrical shape, and then transmit its precise size to digital image. The precision of this method is subject to the standard part's accuracy of manufacture and the picture border's inspection quality. Moreover, it has strict requirements on the shape of standard part. In addition, the standard part's border must be parallel to the direction of CCD pixel.

Or the distance between parallel lines will correspond to slant image's pixel number. And there will be angular misalignment in the demarcated pixel equivalent, as the picture 1 (b) shows. To optimize the calibration method, this paper presents an improved SIFT-based method to demarcate the pixel equivalent of Microscopic inspection system, according to its calibration characteristics. SIFT-Feature point matching method is based on image's local gradient features. It has better adaptability to image's lighting change, fuzziness and distortion. It doesn't need to detect the object's border. And it's not sensitive to object's shape.



(b)

Fig.1 Traditional method of calibration

2 Calibration Technique

Micro-vision system, which is based on SIFT- algorithm, can accomplish the calibration process with only the measured object. It doesn't have to draw support from standard part, which changes the traditional method that acquires the number of pixels of the reference distance. Detail calibration steps as follows: 1) Lay the object upon the one-dimensional translation stage; 2) Move the object. And collect the images of it separately before and after the moving; 3) Read the object's displacement L through the numerical reading device (optical grating or other displacement output unit); 4) Extract SIFT characteristic points of the collected images; 5) Match SIFT characteristic points; 6) calculate the number of pixel's corresponding distance the characteristic points moved; 7) L/N is the pixel equivalent of the vision system.

What is key in the calibration method base on SIFT algorithm is the extraction and matching of SIFT characteristic points.

2.1 Method of Abstracting SIFT Characteristic Points

The SIFT algorithm extracts extreme points of Difference of Gaussian scale-space(DoG) as characteristic points, which mainly contains the following 4 steps[3-8]:

1) Build a DOG and extracts its extreme points;

$$\begin{aligned} D(x,y,\sigma) &= (G(x,y,k\sigma)-G(x,y,\sigma))*I(x,y) \\ &= L(x,y,k\sigma)-L(x,y,\sigma) \end{aligned} \tag{1}$$

$G(x,y,\sigma)=\frac{1}{2\pi\sigma^2}e^{-(x^2+y^2)/2\sigma^2}$ is a scale variable Gaussian function. *L(x,y,σ)* is an Gaussian image. k is a scale scaling factor. σ is a Gaussian scale. The extreme point is the one that has the biggest change in grayscale value compared with its surrounding 26 pixels.

2) Precisely position extreme value points through fitting 3-d quadratic function.

3) Direction distribution of characteristic points

Calculate the images' gradient of characteristic points in the DoG. Determine the characteristic points' main direction by the way of accumulating their gradient. The amplitude and direction angle of gradient are calculated as follows:

$$\begin{cases} m(x,y)=\begin{bmatrix}(L(x+1,y)-L(x-1,y))^2 \\ +(L(x,y+1)-L(x,y-1))^2\end{bmatrix}^{1/2} \\ \theta(x,y)=\tan^{-1}\frac{L(x,y+1)-L(x,y-1)}{L(x+1,y)-L(x-1,y)} \end{cases} \tag{2}$$

4) Build descriptor of the characteristic point

Turn the coordinate axis of Gaussian image *L(x,y,σ)* to the main direction of characteristic points. The characteristic points being the center, divide its surrounding 16×16 pixel window into 4×4 windows. In each small window, accumulate the Gaussian image's gradient projection in 8 directions. Then a characteristic vector F of 4×4×8=128 dimensions can be obtained.

2.2 Characteristic Points Matching and Mismatch Filtering

After detecting the SIFT characteristic points of the object's image, the points obtained after moving need to be matched with those before moving, thus confirm their one-to-one correspondence. The SIFT characteristic point matching methods that are in common used mainly include: NN(Nearest Neighbor), Exhaustive method, K-D(K-Discriminator Tree), BBF(Best-Bin-First) and so on. Above methods are all based on traditional SIFT characteristic point descriptor, all getting too much error matching.

In the calibration process of micro-vision system, even a few mismatches can bring great deviation to the result. In order to overcome the influence of the mismatches, this paper does a second filtering to the mismatches, using the space coordinate information of characteristic points, which is based on NN(Nearest Neighbor) method. Suppose $K(I_{test})=\{k_1^{test},k_2^{test},..k_M^{test}\}$ is the set of characteristic points of the image before object's moving. And $K(I_{temp})=\{k_1^{temp},k_2^{temp},...k_M^{temp}\}$ is the set of characteristic points of the image after object's moving.

Calculate the nearest and the second nearest distance (*MD_i* and *SMD_i*) between each couple of characteristic points' descriptor (before and after the image's moving). And then calculate the value: SMD_i / MD_i as follows: .

$$MD_i=\min_{j=1}^{M}(dis(F(k_i^{test}),F(k_j^{temp}))) \tag{3}$$

$$SMD_i = \underset{j=1}{\overset{M}{\mathrm{second\,min}}}(dis(F(k_i^{test}), F(k_j^{temp}))) \tag{4}$$

$$prop_i = \frac{SMD_i}{MD_i} \tag{5}$$

If $prop_i > th_1$, then point k_i^{test} and point k_j^{temp} are matching with each other. th_1 is an appointed threshold value. With this step, usually 80% mismatches can be filtered. The larger th1 is, the less the number of matching couples is, and the better the stability is. But when th_1 is too large, the number of matching point couples decreases so sharply that correct positioning is impossible, and not all mismatches can be eliminated. In practice, in most cases the value of th1 is around 1.5.

After initial characteristic point matching, point pairs can be obtained: $\{(s_1^{test}, s_1^{temp}),\ (s_2^{test}, s_2^{temp}),\ \ldots\ (s_L^{test}, s_L^{temp})\}$, which still contains many mismatches. Because the relative position of characteristic points doesn't change after object's moving, a second matching can filter mismatches with the space coordinates information. Calculate the nearest distance md_i^{test}, md_i^{temp} and the second nearest distance $smd_i^{test}, smd_i^{temp}$ separately within each characteristic point space.

$$md_i^{test} = \min_{j=1, j\neq i}^{L}(dis(s_i^{test}, s_j^{test})) \tag{6}$$

$$smd_i^{test} = \underset{j=1, j\neq i}{\overset{L}{\mathrm{second\,min}}}(dis(s_i^{test}, s_j^{test})) \tag{7}$$

$$md_i^{temp} = \min_{j=1, j\neq i}^{L}(dis(s_i^{temp}, s_j^{temp})) \tag{8}$$

$$smd_i^{temp} = \underset{j=1, j\neq i}{\overset{L}{\mathrm{second\,min}}}(dis(s_i^{temp}, s_j^{temp})) \tag{9}$$

If both $md_i^{test} / md_i^{temp}$ and $smd_i^{test} / smd_i^{temp}$ are smaller than the appointed threshold value th_2 ,with md_i>10pixel ,then this couple of matching points be kept. If not, delete it. The value of th_2 is about 0.9 in common case. As is in picture 2, k_1 and k_3, k_2 and k_4 are all considered correct match before the second filtering. But actually, only k_1 and k_3 are the correct couple. k_2 and k_4 are the mismatches. After the second filtering, only the couple of k_1 and k_3 will be retained.

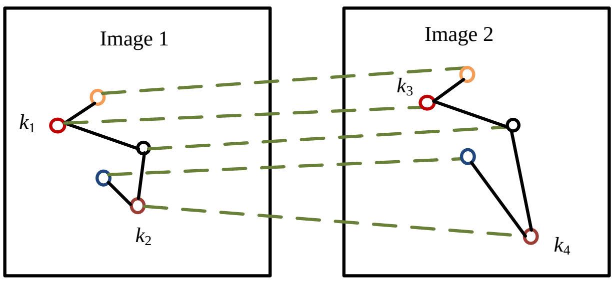


Fig.2 Filter out the mismatches using space coordinates

Lots of experiments show that after the second time filtering, correct correspondence (before and after the object's moving) of object's SIFT characteristic points can be gained. And through affine transformation between images, the number of pixels corresponding to the SIFT characteristic points' moving distance can be calculated. With the reference displacement given, the pixel equivalent of micro-vision system can be demarcated.

3 Experiments and Analysis

In order to examine above ideology, this paper designed two experiments.

3.1 Experiments of Standard Parts' Calibration and Comparison

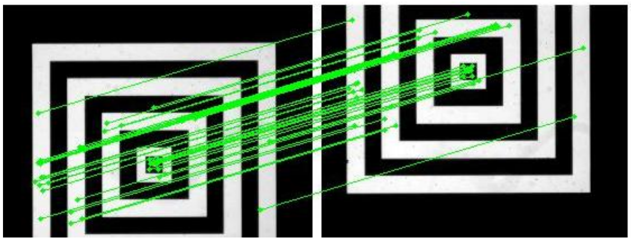


Fig.3 SIFT matching point of the mask

Firstly, lay a standard part (masking layer) upon the one-dimensional electric transition stage which has a grating readout device. Resolution of the optical grating is 0.2μm, using backlit lighting. CCD's type is Japan's Watec 902H, with a nominal pixel size value of 8.6*8.3μm. To make a contrast, pixel equivalent calibration of the standard parts is done separately with the traditional method (including border detection, subpixel subdivision, parallel fitting) and this paper's method. Figure 3 shows the result of SIFT characteristic point matching. And table 1 shows the result of calculation.

Tab.1 Calibration result of pixel using

Number of experiments	1	2	3	4	5
Grating indication (μm)	397.6	494.8	1044.4	1545.4	2192.2
Distance of SIFT points (pixel)	33.6978	41.9511	88.4361	130.7644	185.8332
SIFT pixel equivalent (μm/pixel)	11.7990	11.7947	11.8097	11.8182	11.7966
Base space of mask(μm)	1000.0	2000.0	2000.0	3000.0	3000.0
Measured space of mask (pixel)	84.8	168.7	169.0	254.6	254.5
Pixel equivalent of traditional method (μm/pixel)	11.7925	11.8554	11.8343	11.7832	11.7878

From table 1, the mean value of pixel equivalent is 11.8036 μm/pixel with this paper's method. While with traditional method, the mean value is 11.8106 μm/pixel. Being influenced by border detection's degree of precision, the result of traditional methods has less stability.

In order to check the accuracy of this paper's method further, move the standard part to a certain distance. The displacement can be read out by Renishaw ML10 two-frequency laser; Calculate the displacement's number of pixels with this paper's SIFT algorithm. And then calculate the actual displacement separately with above two calibration results. Table 2 shows the result.

Tab.2 Moving test error of mask image

Number of experiments	1	2	3	4	5
Displacement by Dual-frequency laser (μm)	494.6	504.2	507.4	498.7	495.1
Movement distance of characteristic points (pixel)	41.9270	42.7070	43.0163	42.2561	41.96
Error of pixel equivalent by mine(μm)	0.2895	-0.1037	0.3472	- 0.0741	0.1791
Error of pixel equivalent by traditional method(μm)	0.5830	0.1953	0.6483	0.3699	0.4728

As can be seen in table 2, the measuring result of this method is better than traditional method in application.

3.2 Experiment of Workpiece's Self-Movement

One obvious advantage of this calibration method is that the calibration algorithm is independent from object's shape. So the calibration can be done without standard parts, only making use of workpiece's self-moving. Keeping the same optical magnification as previous experiment, table 3 gives the results of vision system's calibration by workpiece with hole. Two images' matching result of SIFT characteristic points is given in figure 4. In table 3, pixel equivalent error 1 is the deviation of the two calibration results (the one that makes use of workpiece based on this paper's algorithm and the other by traditional method). Pixel equivalent error 2 is the deviation of the two results (the one that makes use of workpiece and the other that makes use of standard part, which all take this paper's algorithm).

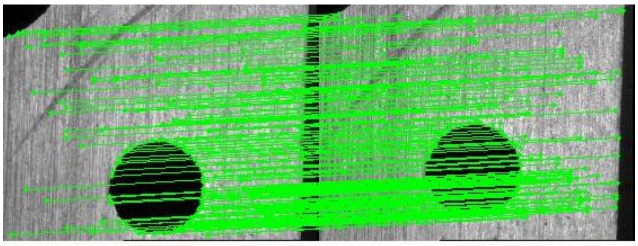


Fig.4 SIFT matching point of the workpiece

Tab.3 Calibration result 1 of work with a hole

Number of experiments	1	2	3	4	5	mean error
Pixel equivalent(μm/pixel)	11.7893	11.8225	11.8105	11.7934	11.8086	
Error 1 of pixel equivalent (μm)	-0.0213	0.0119	-0.0001	-0.0172	-0.002	-0.0057
Error 2 of pixel equivalent (μm)	-0.0143	0.0189	0.0069	-0.0102	0.0050	0.0013

Under different magnification, table 4 gives the calibration result of micro-vision system by workpiece with circular hole. In table 4, pixel equivalent error is the deviation of the two calibration results----the one that makes use of workpiece and the

Tab.4 Calibration result 2 of work with a hole

Number of experiments	1	2	3	4	5	mean error
Pixel equivalent (μm/pixel)	8.7949	8.7997	8.7959	8.8017	8.7892	
Error 1 of pixel equivalent (μm)	0.0052	0.0100	0.0062	0.0120	-.0005	0.0066

one with traditional method.

From table 3 and table 4, we can see that the calibration result by workpiece itself can maintain better repeatability with the method that uses mask. This further again proved this paper's method.

3 Conclusion

In micro-vision measuring system, calibration's precision of pixel equivalent applying traditional methods is subject to standard part. To overcome these shortcomings, this paper proposed a method based on improved SIFT local characteristic point matching to demarcate pixel equivalent. The experimental results show that the method presented in this paper has better stability than the traditional method. While reaching more stable calibration precision, the vision system's positioning accuracy is improved too. With this paper's calibration method, standard part is no more necessary in the process of micro-vision-detecting system's pixel equivalent calibration, overcoming the shortcoming that standard parts need to be positioned precisely.

References

1. Wu Ji-gang, Bin Hong-zan.: Dimensional inspecting system of thin sheet parts based on machine vision. Optics and Precision Engineering, 15, 124. (2007)
2. Wang Jianmin, Pu Zhaobang, Zhao Hui, Lu Zhenzhi.: A new method for calibrating 2-D image measuring systems. Journal of Harbin Institute of Technology, 32, 37(2000).
3. Pan, Xunyu; Lyu, Siwei.: Detecting image region duplication using sift features. Processing of IEEE International Conference on Acoustics, Speech and Signal,1706(2010).
4. Bastanlar. Y, Temizel. A, Yardimci. Y.: Improved SIFT matching for image pairs with scale difference. Electronics Letters, 46, 346(2010).
5. Jing Li, Nigel ,M. Allinson.: A comprehensive review of current local features for computer vision. Neurocomputing,71:1771(2008).
6. David G.Lowe.: Distinctive image features from scale-invariant keypoints . Int. Journal of Computer Vision, 60, 91 (2004).
7. Krystian Mikolajczyk, Cordelia Schmid.: A performance Evaluation of Local Descriptors. IEEE Transactions on Pattern Analysis and Machine Intelligence, 27, 1615(2005)
8.Wu Hui-lan, Liu Guo-dong, Liu Bing-guo, Pu Zhao-bang.: Study on the circle center fast accurate-locating technique based on the SIFT. Journal of Optoelectronics.Laser, 19, 1512 (2008)