http://ansinet.com/itj

ITJ

ISSN 1812-5638

# INFORMATION TECHNOLOGY JOURNAL

ANSInet
Asian Network for Scientific Information
308 Lasani Town, Sargodha Road, Faisalabad - Pakistan

Information Technology Journal 8 (3): 383-387, 2009
ISSN 1812-5638
© 2009 Asian Network for Scientific Information

# Planar Displacement Detection with Point Feature Matching

[1]Chen Feng-Dong, [2]Hong Bing-Rong and [2]Liu Guo-Dong
[1]Department of Automatic Testing and Control, Harbin Institute of Technology, China
[2]School of Computer Science and Technology, Harbin Institute of Technology, China

**Abstract:** A novel planar displacement detection method is implemented using Scale Invariant Feature Transform (SIFT) point feature matching on a calibrated optical grating-vision measuring platform. SIFT is a method for extracting and describing image key-points, which are robustly invariant to scale, rotation and translation as well as robust to illumination changes and limited changes of viewpoint. The platform is moved along its x axis step by step and a series of images are captured with corresponding grating sensor values. SIFT feature points are extracted and matched between the successive images through a K-Dimension Tree (KD-Tree) based feature matching algorithm to detect the displacement of each step. The detected values are compared with the corresponding grating sensor values. Experimental results prove that the accuracy of the method is less than 10 μm in this environment.

**Key words:** Difference of Gaussian, optical grating sensor, scale invariant feature transform, sub-pixel

## INTRODUCTION

As the industries are rapidly developing, the required displacement-detecting level is decreasing. In many inspection and measurement processes, displacement detection plays an important role and the required accuracy and resolution are range from sub-micron level to several hundreds micrometers. There are many displacement detection methods which have different accuracy and can be applied in different conditions (Ding-June *et al*., 2001; Sasaki *et al*., 2000). Image-based displacement detection method has many advantages, such as no touch, fast, low price and rich information (Weng-Bing *et al*., 2004). But it needs image matching which is a hard work especially in the condition of high accuracy. Traditional image matching methods are sensitive to noise, illumination and rotation (Wei-Wen and Hui, 2005). David (2004) came up with the SIFT algorithm which can extract distinctive invariant features from images that can be used to perform reliable matching between different views of an object or scene. SIFT features are invariant to image scale and rotation and are shown to provide robust matching across a substantial range of affine distortion, change in 3D viewpoint, addition of noise and change in illumination. In addition, SIFT features are localized to sub-pixel accuracy. SIFT algorithm gained increasing popularity in object and scene recognition (Wu-Zu *et al*., 2008; Ze-Su *et al*., 2008). This study implements a planar displacement detection method based on SIFT feature matching on a calibrated grating-vision sensor platform.

## SIFT FEATURES EXTRACTING

SIFT algorithm finds key-point at the maxima of a difference-of-Gaussian (DoG) function, which are the most stable image features. SIFT algorithm consists of four major stages: (1) scale-space peak selection; (2) key-point localization; (3) orientation assignment; (4) key-point descriptor.

In the first stage, extreme points are identified by scanning a series of DoG images (Gaussian pyramid) over location and scale. An efficient approach to construction of Gaussian pyramid is as follows:

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

where, I(x, y) is the input image, * is the convolution operation in x and y, σ is a scale factor, k is a const factor and:

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

The input image is incrementally convolved with Gaussians to produce images separated by a constant factor k in scale space. We choose to divide each octave of scale space (i.e., doubling of σ) into an integer number, s, of intervals, so $k = 2^{1/s}$. Adjacent image scales are subtracted to produce the DoG images. Once a complete octave has been processed, we resample the Gaussian image that has twice the initial value of σ by taking every

**Corresponding Author:** Chen Feng-Dong, Department of Automatic Testing and Control, Harbin Institute of Technology, China

second pixel in each row and column. The accuracy of sampling relative to σ is no different than for the previous octave, while computation is greatly reduced.

Each pixel in the DoG images are checked for being local extreme by comparing with the 8 neighboring pixels on the same level of the pyramid and the 18 pixels from the image patches in the immediately adjacent layers.

In the second stage, candidate key-points are localized to sub-pixel accuracy and eliminated if found to be unstable. An accurate position fix on the key-points located in the previous step has been implemented by fitting a 3D quadratic function to the local sample points:

$$D(X) = D + \frac{\partial D^T}{\partial X}X + \frac{1}{2}X^T\frac{\partial^2 D}{\partial X^2}X \quad (3)$$

where, D and its derivatives are evaluated at the sample point and $X = (x, y, \sigma)^T$ is the offset from this point. The location of the extreme, $\hat{X}$, is determined by taking the derivative of this function with respect to x and setting it to zero, giving:

$$\hat{X} = -\frac{\partial^2 D^{-1}}{\partial X^2}\frac{\partial D}{\partial X} \quad (4)$$

The $\hat{X}$ is added to the location of its sample point to get the interpolated estimate for the location of the extreme.

The third stage assigns the dominant orientations for each key-point based on its local image patch. This step makes the descriptor rotation invariant. It involves calculating the gradient vectors in a window around the SIFT feature on the scale at which the feature was detected. So, for example, if a SIFT feature was detected with scale 3 (third level of the DoG pyramid), the third layer of the Gaussian pyramid is accessed and the gradients of the intensity values are calculated:

$$\left.\begin{aligned} u &= L(x+1,y) - L(x-1,y)\\ v &= L(x,y+1) - L(x,y-1)\\ m(x,y) &= \sqrt{u^2+v^2}\\ \theta(x,y) &= \tan^{-1}(u/v) \end{aligned}\right\} \quad (5)$$

where, m(x, y) and θ(x, y) indicate the magnitude and direction of the gradient at location (x, y).L(x, y) is the intensity of the pixel at location (x, y).

The gradient orientations are computed in a 16×16 window around the SIFT feature and quantized in 8 steps of 45 degree intervals in this study as shown in Fig. 1a. A histogram is incremented by a sample that is weighted by its gradient magnitude and a Gaussian-weighted circular kernel that is placed on top of the 16×16 window. This has

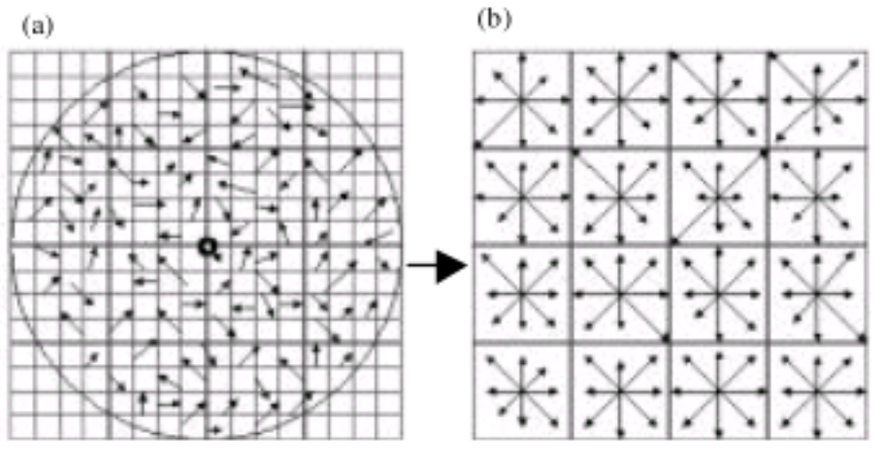


Fig. 1: Gray gradient and SIFT feature vector, (a) Image gradients and (b) Orientation histograms

the effect of giving a higher weight to the samples near the centre of the window. Peaks in the orientation histogram correspond to dominant direction of local gradients and the window is rotated so that the majority gradient is perpendicular to the top margin of the window.

The final step associates each SIFT feature point with a 128-element feature vector that uniquely identifies that point. The gradient orientations obtained at the end of step 3 are arranged into 16 histograms shown in Fig. 1b. Values of the orientation histogram constitute the 128-dimensional vector (8 orientations×16 histograms). This 128-dimensional vector is the SIFT feature vector which is then normalized to unit length.

## KD-TREE BASED FEATURE MATCHING

A KD-Tree based feature matching algorithm is presented for determining the matching SIFT features pairs in adjacent images. The matched pairs are used for the displacement detection. Given a SIFT key-points set E and a target key-point d, then a nearest neighbor of d, d' is defined as:

$$\forall d'' \in E, |d - d'| \le |d - d''| \quad (6)$$

Where:

$$|d - d'| = \sqrt{\sum_{i=1}^{k}(d_i - d'_i)^2}$$

and $d_i$ is the ith dimension of d. The KD-Tree based feature matching algorithm is described as follows: A KD-Tree is a data structure (balanced binary tree) for storing a finite set of points from a k-dimensional space. To build a KD-tree of the key-points set E of the image $I_t$, we start from the root-node and spilt recursively the key-points through their ith dimension. The root-node represents the entire E. The son node is assigned to the lift or right sub-tree if the value of its ith dimension is bigger or smaller than that of father node. For each query

key-point kp of the next image $I_{t+1}$, we find two most nearest neighbors $kp_1$ and $kp_2$ in E. As proved in our experiment, if $|kp_1-kp|/|kp_2-kp|$ is bigger, then the matching quality between kp and $kp_1$ is much higher, otherwise the quality is lower. So, we can use the following equation to judge the matching for two key-points:

$$|kp_1-kp|/|kp_2-kp|<\alpha \quad (7)$$

where, $\alpha$ is constant and $0<\alpha<1$.

If this Eq. 7 is satisfied, the matching is successful. By choosing proper factor $\alpha$, we can eliminate most false matching pairs and get excellent matching performance. In order to find the closest point, the KD-Tree is searched recursively. The query point needs to be compared with the separating plane in order to decide on which side the search must continue. This procedure is executed until the leaves are reached. All bucket points will be evaluated. However, the closest point may be in a different bucket, if the distance to the limits is smaller than the one to the closest point in the current bucket, Backtracking will be performed. Searching for closest points using KD-Tree needs logarithmic time O(log N), i.e., the amount of backtracking is independent of the number of stored points in the tree.

## CCD CAMERA CALIBRATION

The optical grating-vision measuring platform consists of three independent axes. The translation value of each axis can be recorded by three calibrated optical grating sensors whose calibrating accuracy is 1 μm.

The notation adopted here is that (Fig. 2) the world coordinate systems is described as $S_w$:<$O_w$ $X_w$ $Y_w$ $Z_w$>, the camera coordinate systems is described as $S_i$:<$O_i$ $X_i$ $Y_i$ $Z_i$> and the optical grating coordinate systems is described as $S_g$:<$O_g$ $X_g$ $Y_g$ $Z_g$>. Because of the orthogonal nature, the transformation between $S_w$, $S_i$ and $S_g$ can be briefly expressed as follows:

$$\left.\begin{aligned}\Delta X_w = \Delta X_g = \Delta Y_i u_y\\ \Delta Y_w = \Delta Y_g = \Delta X_i u_x\end{aligned}\right\} \quad (8)$$

where, $u_x$ and $u_y$ are the transformation factor of x and y axes between $S_w$ and $S_i$. When an object is put onto the platform, we move the body of CCD camera along the z axis to focus on the object. The lens of the CCD camera is not adjusted. So $u_x$ and $u_y$ are constant in this condition. We use a standard mask plate (Fig. 3) to calibrate $u_x$ and $u_y$. The standard mask plate is a flat glass coated with specified size of rectangular film by a chemic coating method. The accuracy of the width of the rectangular film is less than 1 μm. So, it can be used as a ruler for

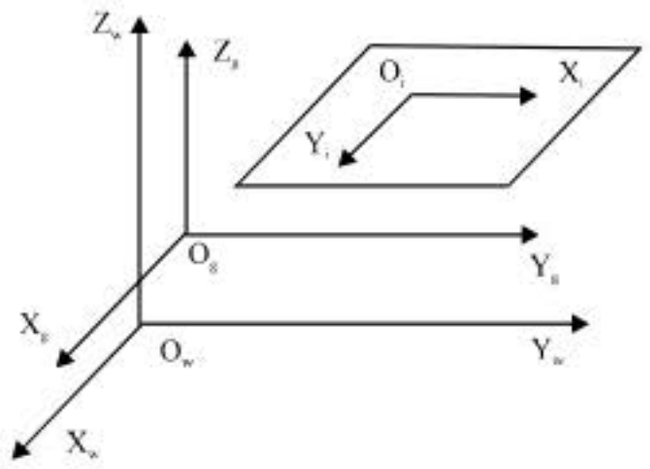


Fig. 2: The coordinate systems of the platform

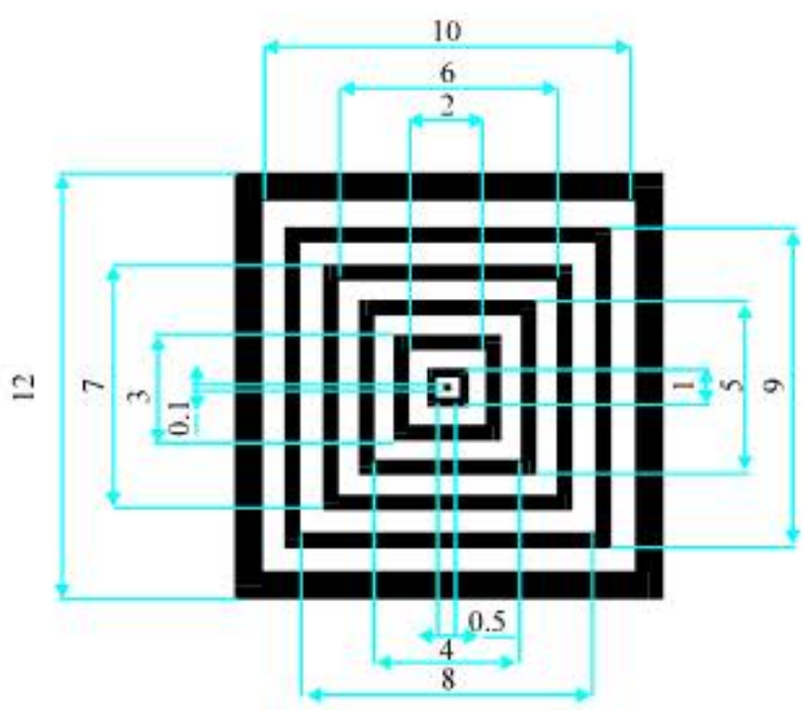


Fig. 3: The standard mask plate

CCD calibration. It is captured the image of the mask plate and extracted the sub-pixel edge of the rectangular film. We got the width of the rectangular film in pixels. So we can calibrate the CCD camera with sub-pixel accuracy. The final result is $u_x = 11.7223$ μm pixel$^{-1}$ and $u_y = 11.4811$ μm pixel$^{-1}$.

## RESULTS AND DISCUSSION

Experiments were performed on the calibrated optical grating-vision measuring platform incorporating a 2.0 GHz Intel Pentium processor. A CCD camera with effective field of view of 768×576 is used for capturing the images. The following experiments were carried out online, i.e., the platform was moved alone its x axis step by step with a step length Δg. A serial of images ($I_t$, $I_t$,...,$I_{t+n}$) were captured with corresponding values of optical grating sensor ($G_t$, $G_{t+1}$,...,$G_{t+n}$). The displacement of each step can be detected by matched SIFT feature points pairs:

$$\Delta X_{t+1} = \mu_y \frac{1}{n}\sum_{i=1}^{n}(x_i^{t+1} - x_i^t) \quad (9)$$

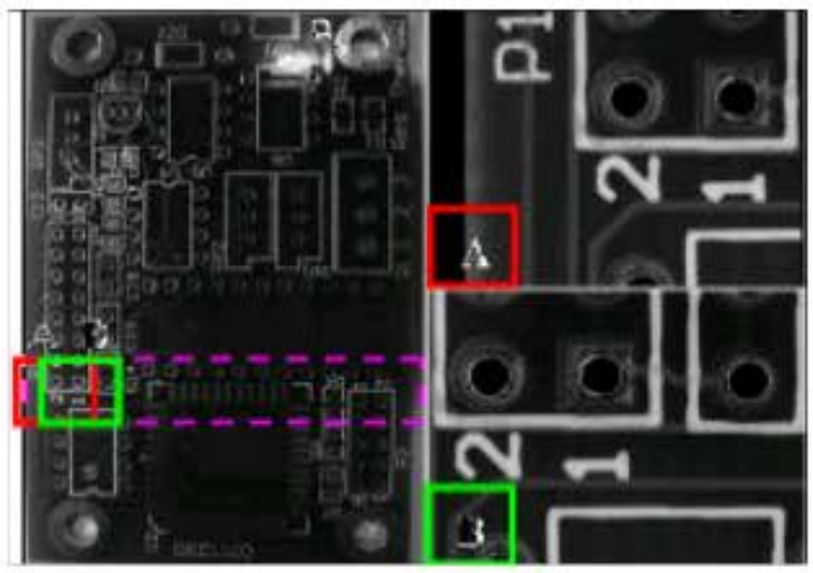


Fig. 4: The PCB image and its first and second images of row 2

where, $x_i^t$ and $x_i^{t+1}$ are the x location of the matched points of image $I_t$ and $I_{t+1}$.

The difference (D-value) of each step between point matching and the optical grating sensor is:

$$D_{t+1} = \Delta G_{t+1} - \Delta X_{t+1} \quad (10)$$

Where:
$\Delta G_{t+1} = G_{t+1} - G_t$

We first captured a PCB whose size is bigger than the field of view of the CCD camera. We captured 48 images (4 rows and 12 images each row). Figure 4 shows the PCB image and the first and second images of row 2 (A and B).

For image matching, SIFT features are first extracted from image and stored in a database. A new adjacent image $I_{t+1}$ is matched by individually comparing each feature from the image $I_{t+1}$ to this earlier database and finding candidate matching features based on their feature vectors. In this study, we directly match adjacent images through KD-Tree based feature matching algorithm to detect the displacement of each moving step. For example, there are 274 and 267 SIFT feature points were extracted from the first and second images of row 2, as shown in Fig. 5a and b. And the 148 matched SIFT features pairs are found as shown in Fig. 5c. The location of the circle points in Fig. 5c equals that of matched points in Fig. 5a and the location of cross mark points equals that of matched points in Fig. 5b. The linking lines between them are the detected displacement. The mean value of the 148 matched features pairs is used as the detected result.

We captured 4 rows and 12 images in each row and all the detected displacement values by points matching are shown in Table 1.The corresponding displacement values recorded by the optical grating sensor are shown in Table 2.

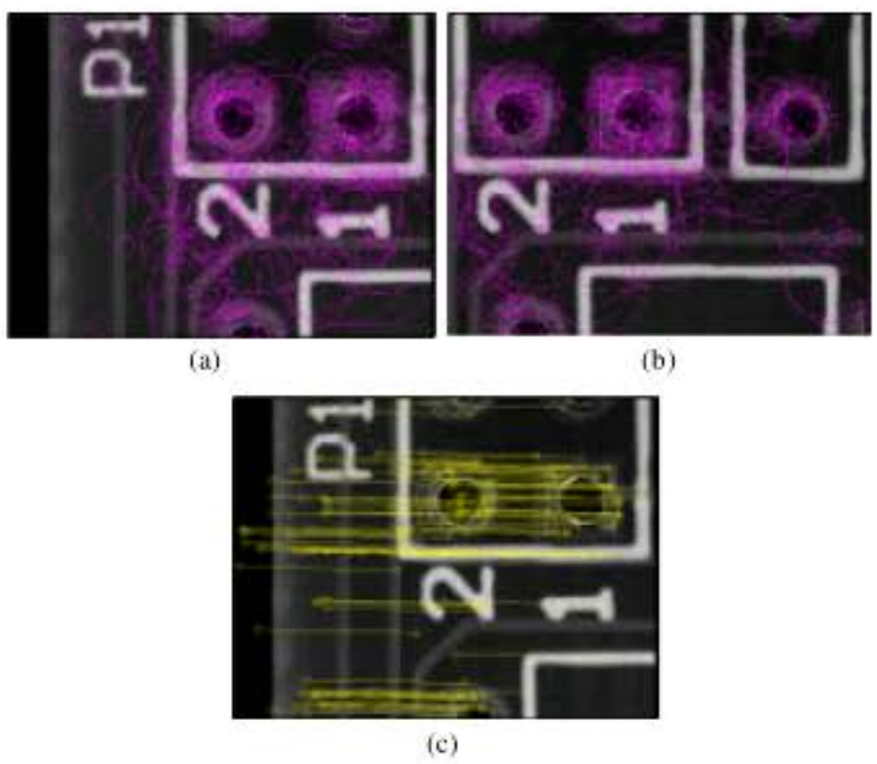


Fig. 5: Feature extracting and matching of the PCB

Table 1: Displacement detected by point matching (mm)

| Displacement value | $Row_1$ | $Row_2$ | $Row_3$ | $Row_4$ |
|---|---|---|---|---|
| $\Delta X_1$ | 3.5353 | 3.4945 | 3.5356 | 3.4989 |
| $\Delta X_2$ | 3.5250 | 3.5031 | 3.5528 | 3.5124 |
| $\Delta X_3$ | 3.5289 | 3.4937 | 3.5343 | 3.4974 |
| $\Delta X_4$ | 3.5350 | 3.5004 | 3.5437 | 3.5123 |
| $\Delta X_5$ | 3.5179 | 3.5066 | 3.5262 | 3.5107 |
| $\Delta X_6$ | 3.5132 | 3.4922 | 3.5090 | 3.4924 |
| $\Delta X_7$ | 3.4992 | 3.4951 | 3.5086 | 3.4944 |
| $\Delta X_8$ | 3.4824 | 3.4966 | 3.4790 | 3.4904 |
| $\Delta X_9$ | 3.4794 | 3.4975 | 3.4621 | 3.5001 |
| $\Delta X_{10}$ | 3.4250 | 3.4928 | 3.4147 | 3.4951 |
| $\Delta X_{11}$ | 3.4705 | 3.5024 | 3.3721 | 3.4924 |

Table 2: Displacement detected by optical grating sensor (mm)

| Displacement value | $Row_1$ | $Row_2$ | $Row_3$ | $Row_4$ |
|---|---|---|---|---|
| $\Delta G_1$ | 3.5376 | 3.4896 | 3.5378 | 3.4918 |
| $\Delta G_2$ | 3.5318 | 3.5008 | 3.5488 | 3.5040 |
| $\Delta G_3$ | 3.5252 | 3.4946 | 3.5314 | 3.4970 |
| $\Delta G_4$ | 3.5308 | 3.5022 | 3.5370 | 3.5040 |
| $\Delta G_5$ | 3.5118 | 3.5042 | 3.5266 | 3.5060 |
| $\Delta G_6$ | 3.5104 | 3.4878 | 3.5056 | 3.4904 |
| $\Delta G_7$ | 3.5040 | 3.4918 | 3.5036 | 3.4924 |
| $\Delta G_8$ | 3.4812 | 3.4876 | 3.4802 | 3.4890 |
| $\Delta G_9$ | 3.4768 | 3.4980 | 3.4706 | 3.4960 |
| $\Delta G_{10}$ | 3.4202 | 3.4916 | 3.4104 | 3.4918 |
| $\Delta G_{11}$ | 3.4770 | 3.4944 | 3.3690 | 3.4944 |

Figure 6 is the D-value of the 4 rows between point matching and the optical grating sensor which prove that the displacement measured by image matching is less than 10 μm.

Another experiment was performed with the same experiment condition using a different object: a circular target shown in Fig. 7. The D-value between point matching and the optical grating sensor is shown in Fig. 8. This experimental result also proves that the displacement measured by image matching is less than 10 μm.

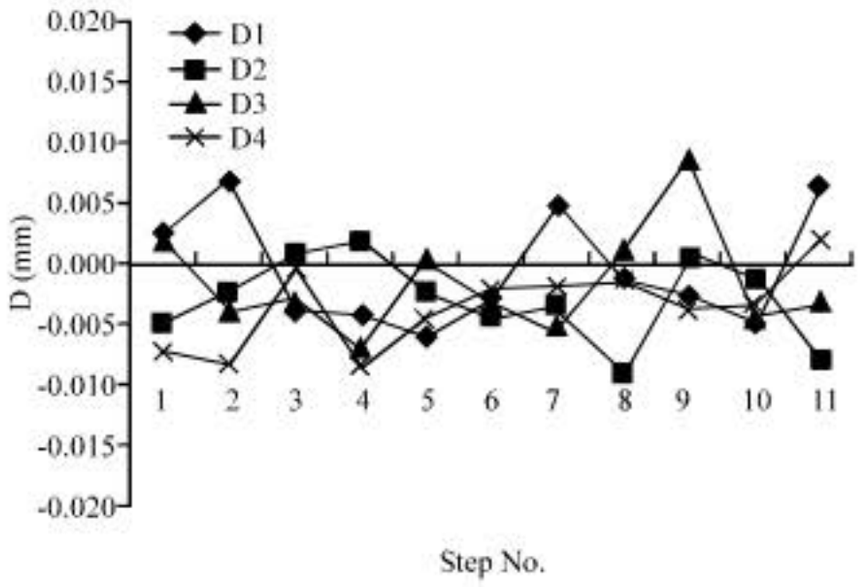


Fig. 6: D-value between point matching and the optical grating sensor of the PCB

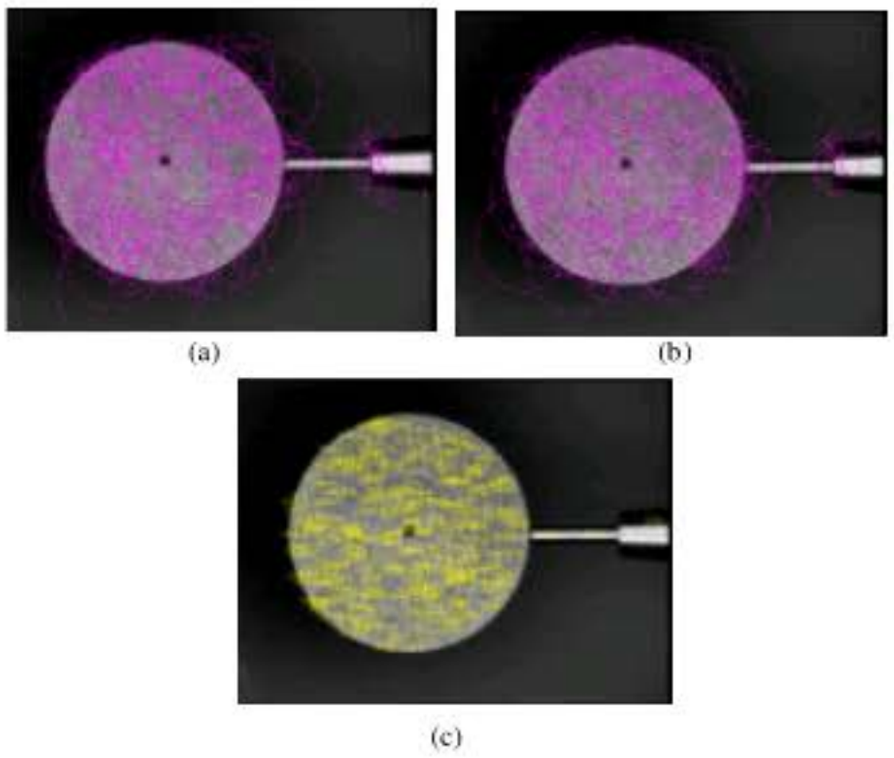


Fig. 7: Feature extracting and matching of the circular target

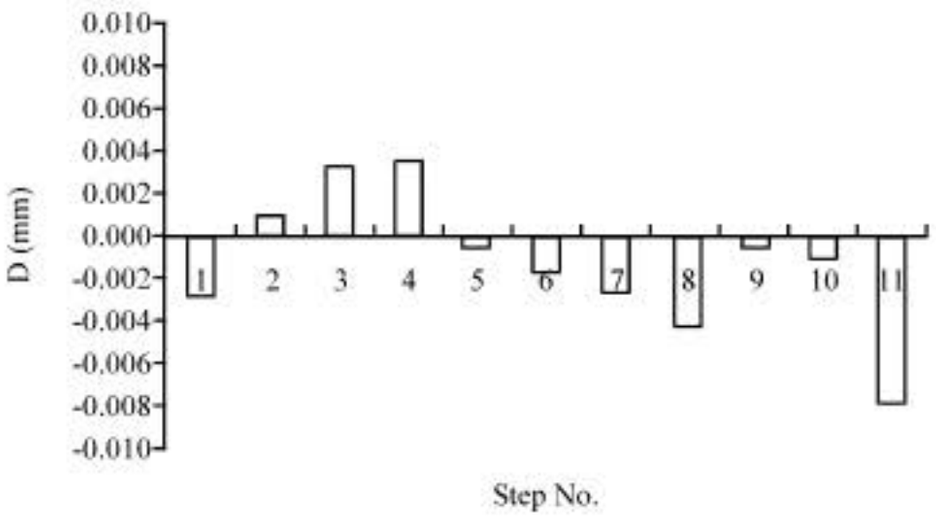


Fig. 8: D-value between point matching and the optical grating sensor of the circular target

## CONCLUSION

A novel planar displacement detection method is implemented using SIFT point features on a calibrated optical grating-vision measuring platform in this study. The CCD camera was calibrated by a standard mask plate. A KD-Tree based feature matching algorithm is implemented which can eliminate most false matching pairs and bring excellent matching performance. Two experiments were carried out and the results showed that this method has the advantages of high stability and high accuracy less than 10 μm.

## REFERENCES

David, G.L., 2004. Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vision, 60: 91-110.

Dong-June, C., C. In-Mook and S.H. Kim, 2001. Inductive micro displacement detecting system with high sensitivity and low linearity error. Int. J. Korean Soc. Precision Eng., 2: 54-60.

Sasaki, M., T. Hirano and X. Mi, 2000. Short-range displacement detection from speckle interference using transparent thin-film photodiode. J. Opt. A: Pure Applied Opt., 2: 534-537.

Wei-Wen, L. and Z. Hui, 2005. Microscopic displacement detection method based on the imaging technology and its application in stiffness measurement. Chinese J. Sensors Actuators, 18: 385-387.

Wen-Bing, L., F. Ping and C. Zeng-Sen, 2004. Displacement measuring based on digital image processing. J. Zhejiang Univ. Technol., 6: 112-118.

Wu-Zu, Y., H. Huang-Xin, D. Li-Xin, M. Wang and C. Yan-Huai, 2008. A simultaneous localization and mapping method based on fast-hough transform. Inform. Technol. J., 7: 190-194.

Ze-Su, C., C. Jian, L. Sun and M. Li., 2008. Novel RBPF for mobile robot SLAM using stereo vision. J. Artificial Intell., 1: 1-11.