

# ROBUST REGISTRATION OF DEFECT SET BY LOCAL CONSISTENCY OF POINT DATA

Takahiro Urano, Shun'ichi Kaneko, and Takayuki Tanaka

Graduate School of Information Science and Technology, Hokkaido University, Kita-14, Nishi-9, Kita-ku, Sapporo, 060-0814, JAPAN;

Commission V/1

**KEY WORDS:** Registration, defect inspection, robustness, consistency

## ABSTRACT:

A novel registration algorithm using Local consistency of point dispersion (LCPD) is proposed, which is based on a new type of measure of consistency between two sets of point data. It is expected to be effective and robust in registration of randomly distributed inconsistent observed data points, and has been recognized to bear wide range of convergence into true solutions through experiments with real world data.

## 1 INTRODUCTIONS

Recently, there has been a rapid and large development in miniaturization and high capacity in the semi-conductor products, such as flash memories, and the memory device media, such as HD and DVD(The International Technology Roadmap for Semiconductors: 2005, 2005). Since, in the production lines for these products, many kinds of microscopic defects can cause functional difficulty, therefore it is requested to develop effective inspection methods which can handle the above-mentioned difficulties. Because the occurrence and characteristics of these defects strongly relates to their production processes, it is possible to infer causes of defects and to control yield rates through the observation and investigation of characteristics of defect distributions representing their positions and sorts. Furthermore, since many inspection systems with high sensitivity and performance have enabled to detect a lot of defects from even simple media, effective and efficient methods of stability are required to process the set of many defects with high precisions even quickly. Additionally, as one of the important pre-processing or tasks that should be done before inspection, rotational registration is necessary for analysis and classification of the defects in the cases of the disk-like products without no physical indices like notches, such as HD and DVD, that is very different from semi-conductor wafers with the notches for registration. In the inspection stages in practical production lines, in order to handle lots of requirements on advanced inspection, not only system integration with heterogeneous sensors has been popular but also sensors in different environmental conditions should be utilized to obtain combinational observations and to analyze some sorts of the defects. For realizing these advanced systems, sets of the observed data points should be registered each other, which may include peculiarities and biases and furthermore the total numbers of data or local relative positions may be different. In order to manage these point data with local inconsistencies in rotational registration, we have to also consider their characteristics of uniform distribution not in any particular regions.

As a simple method for registration of point data, Iterative closest point algorithm (ICP for short) has been well known in this field and well analyzed and applied to the real problems. In ICP, two sets of point data are iteratively merged to obtain as a solution of the least square evaluation of a summation of residuals in three dimensional position. There have been some extensions of the original version of ICP by introducing robust statistics-based functions, such as M-ICP (S.Kaneko et al., 2003, H.Okuda et al.,

n.d.) and the other (), which aims to obtain coincidence between two data sets with non-overlapping regions, however, the local inconsistency above-mentioned are not thought as easy problems to solve effectively.

Another point to be considered in this study is how to evaluate matching between point sets. Almost all the propositions have adopted minimization of summation of point-wise absolute difference or residual, sometimes in their squared values. But in the above-mentioned situations or conditions, possibly the same defect on a media surface can be often measured in a little different positions by different types of sensors, therefore we should try to design some method which can treat these irregular point data in different and reasonable evaluation schema. Furthermore, as in the extended approaches of ICP, consistent registration of the data with inconsistencies by occlusion and addition or subtraction of point data has to be realized. On iterative computation for convergence into global minimum in the search space, possibly wider range around initial solutions should be ensured in order to promote many applications of the propositions.

In this study, in order to solve the above-mentioned problems, we propose a novel registration called Local consistency of point dispersion (LCPD for short), which is based on a new type of measure of consistency between two sets of point data. In the proposed method, for evaluation of the consistency in local regions around the points, we introduce counting *color* or belonging of neighboring points, which represents how sparse or crowded it is. The proposed method is expected to be effective and robust in registration of randomly distributed inconsistent observed data points, and has been recognized to bear wide range of convergence into true solutions through experiments with real world data.

## 2 LOCAL CONSISTENCY OF POINT DISPERSION (LCPD)

The proposed method for registration of point data is based on a novel type of measure of consistency in local distributions between two sets of point data, which is named as Local consistency of point dispersion, hereafter LCPD for short. Figure 1 shows the schematic procedures of the proposed method. Between two sets of point data are evaluated by LCPD in Figure 1(1). The evaluation of global inconsistency is measured by gathering elements, and then it becomes high if two sets of point data are coincident

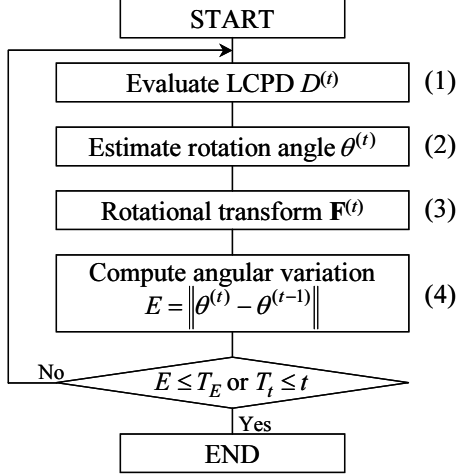


Figure 1: The schematic procedures

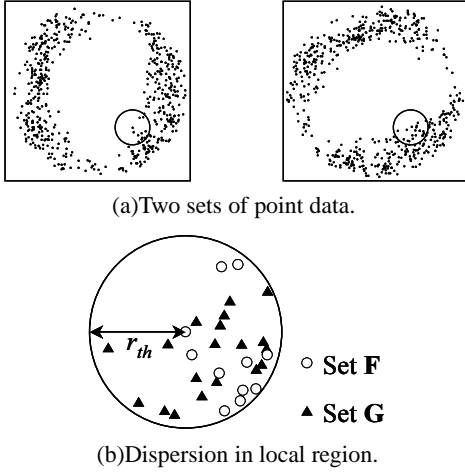


Figure 2: Inconsistency in a local region.

in density. By using LCPD, we can realize robustness in registration in the case that the two sets have a difference in number. Figure 2 shows an example of elemental computation of local inconsistency in every local region centered at each point in the set  $\mathbf{F}$ , where the inconsistency is defined a difference in numbers of the neighbouring points drawn from both the set  $\mathbf{F}$  and  $\mathbf{G}$ . In general, if we have the large difference in number or distribution then the inconsistency is assumed to be high.

We formalize the abovementioned principle of evaluation of LCPD. Define  $\mathbf{F}^{(t)} = [\mathbf{f}_1^{(t)}, \mathbf{f}_2^{(t)}, \dots, \mathbf{f}_M^{(t)}]$  and  $\mathbf{G}^{(t)} = [\mathbf{g}_1^{(t)}, \mathbf{g}_2^{(t)}, \dots, \mathbf{g}_N^{(t)}]$  as two sets of point data by using the data number of  $M$  for  $\mathbf{F}$  and  $N$  for  $\mathbf{G}$ .  $\mathbf{F}$  is iteratively rotated and merged in order to find a match with  $\mathbf{G}$ . A rotationally transformed data  $\mathbf{F}^{(t)}$  at  $t$ -th step is calculated as

$$\mathbf{F}^{(t)} = \mathbf{R}^{(t)} \cdot \mathbf{F}^{(t-1)}, \quad (1)$$

where  $\mathbf{R}^{(t)}$  is the rotation matrix for the sequential transformation at the  $t$ -th step and it is defined as

$$\mathbf{R}^{(t)} = \begin{bmatrix} \sin \theta^{(t)} & \cos \theta^{(t)} \\ \cos \theta^{(t)} & -\sin \theta^{(t)} \end{bmatrix} \quad (2)$$

where  $\theta^{(t)}$  shows the rotation angle. Next, a neighbouring set of points  $\Omega_m$  which is defined as a local disk region around each

point  $f_m$  in  $\mathbf{F}$  is given by

$$\Omega_m = \{x \mid \|x - f_m\| \leq r_{th}\}, \quad (3)$$

where a parameter  $r_{th}$  is the radius of the circle region. An elemental definition of local inconsistency  $d$  is defined as

$$d_m^{(t)} = \left| \sum_{\Omega_m} w^f \cdot c_m^f - \sum_{\Omega_m} w^g \cdot c_m^g \right|, \quad (4)$$

where  $c_m^f$  and  $c_m^g$  are the indicating value both in the neighbouring sets defined in  $\mathbf{F}$  and  $\mathbf{G}$  and given by

$$c_m^f = \begin{cases} 1 & : f_i \in \Omega_m \\ 0 & : \text{otherwise,} \end{cases} \quad (5)$$

$$c_m^g = \begin{cases} 1 & : g_n \in \Omega_m \\ 0 & : \text{otherwise.} \end{cases} \quad (6)$$

For example in the case of Fig 2(b), we find the cardinalities  $\sum c_m^f = 11$ , and  $\sum c_m^g = 18$ . Then  $w^f$  and  $w^g$  are the weights which are determined based on a condition like a total number of points. Here, in this paper, we set the weights as  $w^f = 1/M$  and  $w^g = 1/N$  for making balance in cases that data sets include much unbalanced numbers of point data. We define the accumulated LCPD  $D$  as a summed or averaged elemental evaluation defined over the total point data

$$D^{(t)} = \frac{1}{M} \sum_m d_m^{(t)} \quad (7)$$

A solution of the rotation matrix  $\mathbf{R}$  is obtained by optimizing or minimizing the abovementioned evaluation value  $D$  with respect to the rotation angle  $\theta$ .

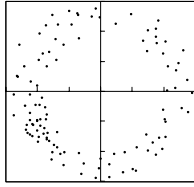
We perform the minimization of the accumulated LCPD  $D$  by use of the descending simplex method(?). The method has been known to be not so efficient but easy in computation because of no differentiation. A simplex is a range with two vertices or boundaries in  $\theta$  because the LCPD is a one dimension function which has only the variable  $\theta^{(t)}$ . The actions of reflection, expansion, reduction and contraction are prepared for the simplex to scan over the search space. For terminating iterations, we calculate an variation  $E$  in angle after  $\mathbf{F}$  is renewed from the step  $t - 1$  to the step  $t$ .

$$E = |\theta^{(t)} - \theta^{(t-1)}| \leq T_E \quad (8)$$

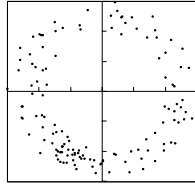
### 3 CONVERGENCE RANGE

A convergence range is defined as the range in search space, which represents a possible range of convergence into true solutions. Since an optional parameter showing a radius  $r_{th}$  is used in the LCPD method, we verify an influence on the convergence range by the radius. Two sets of point data  $\mathbf{F}$  and  $\mathbf{G}$  for this verification are shown in Figure 3(a) and (b). The set  $\mathbf{F}$  has  $M = 117$  points and the set  $\mathbf{G}$  has  $N = 132$  points, and the true solution is  $30[\text{deg}]$ . We compute the evaluated values in LCPD when  $\mathbf{F}$  is turned by every  $0.1[\text{deg}]$  for observation. We change the radius  $r_{th}$ , the profiles of  $D$  are shown in Fig 4(a)through (d) For comparison, ICP(P.J.Besl and N.D.McKay, 1992) was adopted and the profile of the evaluations was shown in the figure. Then results of the convergence ranges by using LCPD and ICP methods are shown in Table 1.

When we use LCPD, the convergence range is around seven times expanded in comparison with ICP regardless of the value of  $r_{th}$ , although it might be dependent on the distribution of points in the sets.



(a) Set F(117points) .



(b) Set G(132points).

Figure 3: Sets of point data

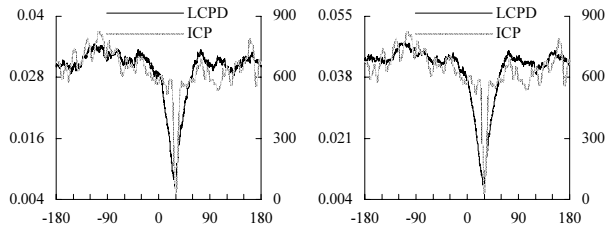
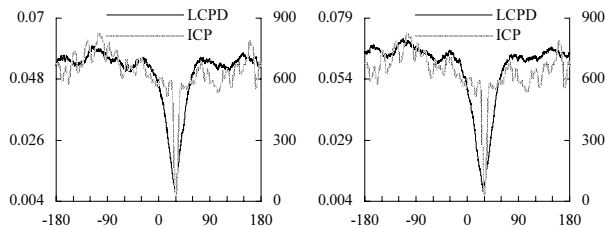
(a)  $r_{th} = 4$ .(b)  $r_{th} = 5$ .(c)  $r_{th} = 6$ .(d)  $r_{th} = 7$ .

Figure 4: The distribution of evaluated value of Figure 3

## 4 EXPERIMENTS

For the experimental verification of the proposed method, we have made many simulated sets of point data and the real data obtained in the real-world inspection system. If not specified, the parameters used in the experiments are  $r_{th} = 5$ ,  $T_E = 0.001$ ,  $T_t = 50$ .

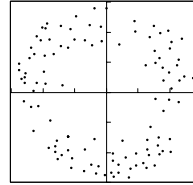
We have performed the comparison between LCPD and ICP with the same simulation data. The simulation data have been made by combination of operations which are shown as follows:

1. Add some points (Ratio: 200%, 500%, 1000%).
2. Move some points (Ratio: 0%, 20%, 50%).
3. Rotate set of points (0 through 60, every 5[deg]).

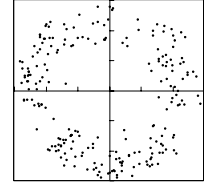
We made 100 sets of point data for each combination. Adding and moving operations are done only according to Gaussian distributions centered at original points in the data. An original point data is shown in Figure 5(a), and Figure 5(b) through (d) show some examples of the sets which were made from the data shown in Figure 5(a). We have performed experiments in order to compare success rates of convergence for both LCPD and ICP based on the simulated point data. We compute a success rate of matching a reference data (Figure 5(a)) with object data which are made by the combination of operations. If the difference in angle of the true angle and the convergence angle is less than  $\pm 1[deg]$ , then the matching was assumed as success. Figure 6 shows the rate of success or convergence for every sets of data. The number in the

Table 1: Convergence ranges of Figure 3

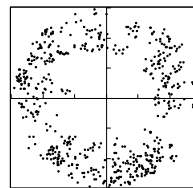
$r_{th}$	LCPD [deg]	ICP [deg]
4	[-60.5, 45.5]	[-8.7, 9.2]
5	[-59.3, 48.1]	
6	[-60.4, 50.5]	
7	[-62.6, 55.2]	



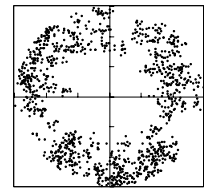
(a) A reference data (113 points).



(b) An object data (226 points).



(c) An object data (565 points).



(d) An object data (1130 points).

Figure 5: Sets of point data.

legends are the ratio of movement in the operation 2. Figure 6 showed the success rates by LCPD, where the convergence performance seems to a little increase as the number of points in **G** increases. We expect the increment in the points or density in **G** may expand handholds which can operate just as pitons on cliffs.

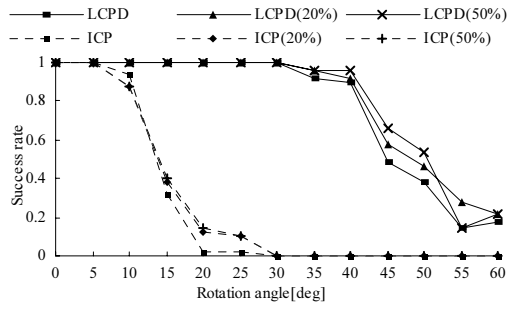
## 5 CONCLUSIONS

We have proposed a novel method of registration named Local consistency of point dispersion (LCPD), which is based on a new type of measure of inconsistency between two sets of point data. The proposed method is expected to be effective and robust in registration of randomly distributed inconsistent observed data points, and has been recognized to bear wide range of convergence into true solutions through the experiments with real world data.

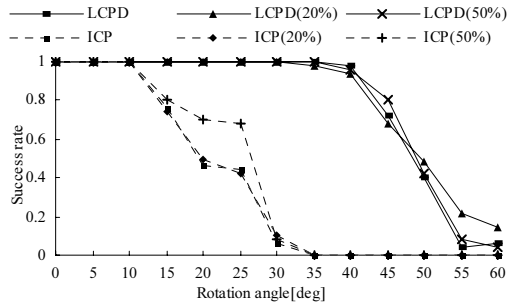
As future works, we will intend to make a method of setting initial conditions in angle which may promise to converge to the true angles. We will consider how to use shifting and matching of histograms(?) which is made from a density of points. Furthermore, we have used all of the data which are the same kind of defect, but there will be really some distinct kinds of defective points even like scratches and foreign substances. So we should make an extended version of the proposed method which can take the characteristics of some kinds of defects into account.

## REFERENCES

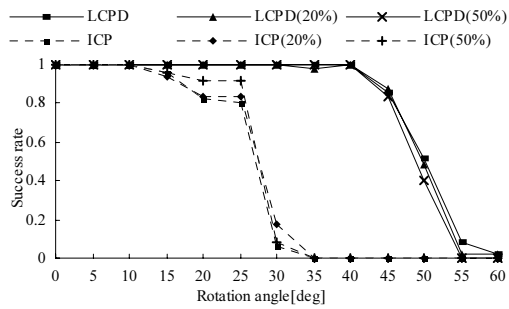
- H.Okuda, Y.Kitaaki, M.Hashimoto and S.Kaneko, n.d. Fast 3-d registration algorithm with hierarchical and region selection approach of m-icp. Journal of Robotics and Mechatronics(In Press).
- The International Technology Roadmap for Semiconductors: 2005, 2005.
- P.J.Besl and N.D.McKay, 1992. A method for registration of 3-d shapes. IEEE Trans. on PAMI 14(2), pp. 673–689.



(a) Matching with the data of 226 points.



(b) Matching with the data of 565 points.



(c) Matching with the data of 1130 points.

Figure 6: Success rate.

S.Kaneko, T.Kondo and A.Miyamoto, 2003. Robust matching of 3d contours using iterative closest point algorithm improved by m-estimation. Pattern Recognition 36(9), pp. 2041–2047.