

# Use of Paraplanar Constraint for Parallel Inspection of Wafer Bump Heights

**Mei Dong**

Dept. of Mechanical and Automation Engineering  
The Chinese University of Hong Kong, Hong Kong, China  
E-mail: {mdong,rchung}@mae.cuhk.edu.hk

**♦Ronald Chung**

**Edmund Y. Lam**

Dept. of Electrical & Electronic Engineering  
The University of Hong Kong, Hong Kong, China

**Kenneth S. M. Fung**

ASM Assembly Automation Ltd.

**Abstract** - *The shrunk dimension of electronic devices leads to more stringent requirement on process control and quality assurance of their fabrication. For instance, direct die-to-die bonding requires placement of solder bumps not on PCB but on the wafer itself. Such wafer solder bumps, which are much miniaturized from the counterparts on PCB, still need to have their heights meet the specification, or else the electrical connection could be compromised, or the dies be crushed, or even the manufacturing equipments be damaged. Yet the tiny size and the texturelessness and mirror nature of the bumps pose great challenge to the 3D inspection process. This paper addresses how a large number of such wafer bumps could in parallel have their heights checked against the specification without the 3D metric information of individual bumps nor their average height be reconstructed explicitly. The system involves two pairs of illumination-plus-camera equipments, and the use of a 3x3 inspection quality matrix that is extracted from the image data. The system is further improved by exploiting the fact that the plane of the bump peaks and plane of the bump bottoms can well be regarded as parallel because the visual field of the inspected die in the image data is small. Experimental results are shown to illustrate the effectiveness of the inspection system.*

**Keywords:** Wafer bump, Height inspection, planar homography, parallel plane-induced homographies

## 1 Introduction

An important goal of the semiconductor industry is to achieve smaller-size chips and circuitry. A direction that has been gaining ground in recent years is the wafer-level packaging (WLP) (also called wafer bumping) technology.

It works by growing miniaturized solder bumps (called wafer bumps) on dies and let the dies on wafer level be bonded together directly through the bumps for electrical connections[1].

Alongside the development of the WLP technology is a demand for improved process control as well as quality assurance. However, because of the much reduced size of the circuitry in WLP, traditional technologies are often not applicable. For instance, the wafer bumps are of diameter only tens of microns. Objects of such size are difficult to have their manufacturing quality inspected with accuracy, yet even tiny errors could be hazardous to the WLP process.

One particular and important inspection needed in WLP is to check whether all the tiny bumps on a wafer are of the same specified height. This is to prevent too loose or too strong contact of the bumps in the bonding process, since too loose contact may cause unreliable electrical connection, while too strong contact may cause damage to the dies or even to the manufacturing equipments. Because of the very small size of the bumps, classical technologies for inspecting PCB solder bumps are not applicable. There have been a few specific methods developed for inspecting wafer bumps by reconstructing them in their entirety in 3D, which is generally an expensive process especially for dimensions of such a miniaturized scale [2][3]. After all, in many applications only the height of the wafer bumps not the entire bump shape is required.

In a previous work [4] we proposed a new visual method that can inspect the height information of a whole array of wafer bumps in parallel without requiring explicit 3D reconstruction. The method requires first a design of the imaging system that lets the image positions of the bump peaks be revealed in two displaced images, so that by the principle of stereo vision 3D information and in particular the height information of the wafer bumps are captured by the image data. The design is depicted in

---

♦ Corresponding Author

Figure 1, which consists of two illumination sources, each of which having its own camera for imaging. One illumination source and the accompanying camera (Source 1 and Camera 1) are collinear in their position and orientation and facing the bump top-down, while the other pair (Source 2 and Camera 2) is configured to face the bump from widely separated positions that are symmetric about the top-down axis. When Source 1 is on, only Camera 1 is on to capture the bump peak which appears as a bright spot in the image due to the mirror-like nature of the bump. When Source 2 is on, only Camera 2 is on to capture the bump peak which again appears as a bright spot in the image due to the symmetry of the system setup.

The method however avoids reconstructing 3D explicitly and instead computes from the image data a measure that is not in metric term. The measure, on comparison with the same measure of the reference or so-called golden wafer, will reveal if the bump heights of the inspected wafer exceed the specification or not.

The underlying idea of the method is to access the plane (which we call the bump plane) in 3D that contains most of the bumps, and examine the separation of this plane from the plane (the substrate plane) that contains the bottom of the bumps. This is depicted in Figure 2. However, it may not be the case that all bump peaks are contained in the above bump plane and those that do not cannot be overlooked, as illustrated by Figure 3. In a way, the method treats height determination of an array of wafer bumps as consisting of two tasks: (1) **Overall bump height Determination**: the determination of the separation of the bump plane from the substrate plane; (2) **Outlier bump Detection**: the identification of the individual bumps whose peaks do not lie on the above majority-defined bump plane. The first task is about quantifying the height of most of the bumps in a global measure, and the second identified the bumps whose heights are outliers to that global measure.

To avoid explicit 3D reconstruction and make processes like camera calibration etc. unnecessary, in determining the overall bump height we do not work on the bump plane and substrate plane directly. Instead, we examine only the image-to-image mapping named homographies  $\mathbf{H}_b$  and  $\mathbf{H}_s$  that are induced by two planes respectively. Induced by a plane, a homography is a mapping that allows the image position of a feature point in one image be predicted from the image position of the same feature point in the other image, provided that the feature point is contained in the inducing plane. With the two homographies  $\mathbf{H}_b$  and  $\mathbf{H}_s$  extracted [11], a measure regarding the separation of the bump plane and substrate plane in 3D is accessible as 3x3 matrix that can be referred as the *biplanar disparity matrix (BDM)*. This matrix is not in metric units, and does not require camera calibration nor explicit 3D measurement. The matrix, on comparison with the corresponding matrix of the reference wafer that

has the right bump heights, reveals whether the inspected wafer deviated too much from the specified bump height or not. As for the outlier bumps, they are detected as those bump peaks whose image positions do not agree with the mapping under the aforementioned bump-plane homography  $\mathbf{H}_b$ . Details of the method are reported in [4].

However, the biplanar-disparity based method has one difficulty. It demands the estimation of two homographies  $\mathbf{H}_b$  and  $\mathbf{H}_s$  from two sets of image point correspondences projected from the two planes respectively. Homography estimation is known to be sensitive to resolution error of the observed image positions and other noise [5], and thus demands more image point correspondences to work with. Yet the distinct point correspondences in the image data that allow the two homographies to be estimated may not be in large number. The estimation errors so incurred will propagate and influence the measure on the overall bump height, and is an important notion.

In this paper, we describe an alternative method that could enhance the accuracy of estimating the two homographies and in turn the overall bump height. The method exploits the fact that, due to the very small size of the die as compared with the typical imaging distance, the visual fields of the bump plane and substrate plane are tiny, and thus the two planes can well be approximated as parallel at least on the portions they are imaged. We make use of this fact to let the two sets of image point correspondences, which are for the estimation of  $\mathbf{H}_b$  and  $\mathbf{H}_s$  respectively, be combined into one, and estimate the two homographies simultaneously with the constraints that they are induced by planes that are parallel. We refer to the 3x3 bump height matrix so resulted as the *paraplanar disparity matrix (PDM)*, and the revised method the *paraplanar disparity method*. Experimental results show that there is significant improvement in the accuracy of the overall bump height that the revised method determines.

The organization of this paper is as follows. We first give a brief review of the original Biplanar Disparity Method. How the biplanar-disparity matrix is constructed and how the two tasks – the overall bump height determination and the Outlier bump Detection – are conducted will be briefed. Then the constraint about the bump plane and the substrate plane being parallel is introduced. We show how it allows the two separates sets of image correspondences – one for the bump-plane homography and the other for the substrate-plane homography – be integrated and the estimation of both homographies be improved. Synthetic and real image data experiments are presented to illustrate the improved performance of the new paraplanar-disparity method.

## 2 Methodology

### 2.1 Review of the Biplanar-disparity Method

In the Biplanar-disparity method, the planar homography matrices are employed as the global parameters to describe the bump plane and the substrate plane respectively.

The planar homography induced by any plane  $\Pi$  in 3D to two images (Images 1 and 2) is determinable up to a scale by the 3D position of the inducing plane  $\Pi$  plus the intrinsic and extrinsic parameters of the two related cameras (Cameras 1 and 2). Suppose we are to represent a plane in 3D by its surface normal  $\hat{\mathbf{n}}$  (whose norm is normalized to unity) and its distance  $d$  to the origin of the reference coordinate system, with the reference coordinate frame chosen as the camera coordinate frame of Camera 1. Then for the bump plane  $\Pi_b$  and substrate plane  $\Pi_s$ , it can be shown [4] that the associated planar homographies  $\mathbf{H}_b$  and  $\mathbf{H}_s$  induced by them to Images 1 and 2 are expressible as:

$$\mathbf{H}_b \cong \mathbf{A}'\mathbf{R}^{-1}(\mathbf{I} - \mathbf{t} \frac{\hat{\mathbf{n}}_b^T}{d_b})\mathbf{A}^{-1} \quad (1)$$

$$\mathbf{H}_s \cong \mathbf{A}'\mathbf{R}^{-1}(\mathbf{I} - \mathbf{t} \frac{\hat{\mathbf{n}}_s^T}{d_s})\mathbf{A}^{-1} \quad (2)$$

where  $\mathbf{A}$ ,  $\mathbf{A}'$ ,  $\mathbf{R}$  and  $\mathbf{t}$  are the intrinsic and extrinsic camera parameters of the two cameras (like the pair shown in Figure 1),  $\mathbf{I}$  is the identity matrix,  $\hat{\mathbf{n}}_b$  and  $\hat{\mathbf{n}}_s$  are the unit surface normals of the planes  $\Pi_b$  and  $\Pi_s$ , and  $d_b$  and  $d_s$  are the distances of the planes from the camera center of Camera 1.

The homographies are then estimated from the correspondences [14] over the bump plane and substrate plane respectively. On this, SVD [8] [13] and robust estimation method [9] such as RANSAC [7] can be used to get best-fit homography to the data, and the bumps whose heights deviate too much from the average height are identified as the outliers to the process.

With the two homographies estimated, a 3×3 matrix called the Biplanar-disparity matrix  $\mathbf{D}$  is defined as

$$\mathbf{D} = \frac{\hat{\mathbf{H}}_b}{\frac{w_b}{\alpha}} - \frac{\hat{\mathbf{H}}_s}{\frac{w_s}{\alpha}} \quad (3)$$

where the two scalars  $\frac{w_b}{\alpha}$ ,  $\frac{w_s}{\alpha}$  can be determined from the partial intrinsic parameters  $\mathbf{A}$  and the partial extrinsic parameters (inter-camera translation  $\mathbf{t}$  and fundamental matrix  $\hat{\mathbf{F}}$ ) of the imaging system using the following:

$$\begin{aligned} \hat{\mathbf{F}}\hat{\mathbf{H}}_s &= \frac{w_s}{\alpha}(\mathbf{A}^{-1})^T[\mathbf{t}]_{\times}\mathbf{A}^{-1} \\ \hat{\mathbf{F}}\hat{\mathbf{H}}_b &= \frac{w_b}{\alpha}(\mathbf{A}^{-1})^T[\mathbf{t}]_{\times}\mathbf{A}^{-1} \end{aligned} \quad (4)$$

It can be shown [4] that the norm of  $\mathbf{D}$  is directly related to the separation of the bump plane and substrate plane. In other words, with certain knowledge (which are global, invariant parameters that can be determined in an off-line process) of the imaging setup, the overall bump height of the array of wafer bumps can be determined without explicit reconstruction of the two involved planes but just by examining the norm of  $\mathbf{D}$ . Such a measure, on comparison with the corresponding measure of the golden wafer, will reveal if the inspected wafer has the overall bump height exceeding the specification or not.

### 2.2 Constraint on Parallel Homographies

Crucial to the accuracy of result in the aforementioned system is the estimation of the homographies  $\mathbf{H}_b$  and  $\mathbf{H}_s$ , yet there are often only limited number of feature points in the image data for that purpose. An important notion is therefore how we can make best use of the available image data for more accurate estimation of the homographies and thereby of the overall bump height.

From practical imaging we observe the following. As the visual field of a die is very small in comparison with the imaging distance, the bump and substrate planes can well be regarded as parallel at least on the segments of the planes that cover the die. In other words, in estimating the two homographies it is natural to enforce the constraint that the inducing planes of the homographies ought to be parallel in 3D. With this, the feature points for the respective estimation of  $\mathbf{H}_b$  and  $\mathbf{H}_s$  can be combined into a single set for the simultaneous estimation of  $\mathbf{H}_b$  and  $\mathbf{H}_s$ . This way, either homography is estimated with a bigger set of image data. Below we elaborate the idea.

In Equations (1) and (2), if the two inducing planes  $\Pi_b$  and  $\Pi_s$  are parallel, we have  $\hat{\mathbf{n}}_b = \hat{\mathbf{n}}_s = \hat{\mathbf{n}}$  and  $d_b = d_s + \bar{d}$ , where  $\bar{d}$  is a scalar that reveals the overall height of the array of bumps. Introducing this parallel-homography constraint into Equations (1) and (2), we have

$$\mathbf{H}_b \cong \mathbf{A}'\mathbf{R}^{-1}(\mathbf{I} - \mathbf{t} \frac{\hat{\mathbf{n}}^T}{d + \bar{d}})\mathbf{A}^{-1} \quad (5)$$

$$\mathbf{H}_s \cong \mathbf{A}'\mathbf{R}^{-1}(\mathbf{I} - \mathbf{t} \frac{\hat{\mathbf{n}}^T}{d})\mathbf{A}^{-1} \quad (6)$$

Since in practical imaging  $\bar{d} \ll d_b$ , we have  $\frac{1}{d+\bar{d}} \approx \frac{1}{d}(1-\frac{\bar{d}}{d})$  to the first order approximation, plus the following:

$$\begin{aligned} \mathbf{H}_b &\cong \mathbf{A}'\mathbf{R}^{-1}\left(\mathbf{I}-\mathbf{t}\frac{\hat{\mathbf{n}}^T}{d_s+\bar{d}}\right)\mathbf{A}^{-1} \\ &\cong \lambda\mathbf{H}_s + \frac{\bar{d}}{d_s}\mathbf{A}'\mathbf{R}^{-1}\mathbf{A}^{-1} - \lambda\frac{\bar{d}}{d_s}\mathbf{H}_s \\ &\cong \lambda\left(1-\frac{\bar{d}}{d_s}\right)\mathbf{H}_s + \frac{\bar{d}}{d_s}\mathbf{A}'\mathbf{R}^{-1}\mathbf{A}^{-1} \end{aligned} \quad (7)$$

for some scaling parameter  $\lambda$ .

Equation (7) can be expressed as:

$$\begin{aligned} \mathbf{H}_b &\cong \mathbf{H}_s + \frac{\lambda'\frac{\bar{d}}{d_s}}{\left(1-\frac{\bar{d}}{d_s}\right)}\mathbf{A}'\mathbf{R}^{-1}\mathbf{A}^{-1} \text{ or more simply} \\ \mathbf{H}_b &\cong \mathbf{H}_s + \lambda\mathbf{A}'\mathbf{R}^{-1}\mathbf{A}^{-1} \end{aligned} \quad (8)$$

for some  $\lambda$  which is a scaling factor related to the camera parameters, the imaging distance, the bump height, and the homography matrix of the substrate plane (because the homography matrix is defined up to an arbitrary scale, the exact matrix we use for each homography in our calculation is one with an arbitrarily fixed norm, say one with unit norm).

What we have here is a precise relationship between two homography matrices induced by two parallel and close planes. With the relationship, we can integrate the correspondences from  $\Pi_b$  and the correspondences from  $\Pi_s$  into a single set of correspondences for simultaneous estimation of the two homographies. Such a process allows more accurate homographies be estimated and in turn more precise 3D measurement be made for height inspection of the wafer bumps.

### 2.3 Estimation of Parallel Homographies

Denoting  $\mathbf{M} = \mathbf{A}'\mathbf{R}^{-1}\mathbf{A}^{-1}$ , from Equation (8) we have  $\mathbf{H}_b$  and  $\mathbf{H}_s$  related by

$$\mathbf{H}_b \cong \mathbf{H}_s + \lambda\mathbf{M} \quad (9)$$

The image correspondences over the bump plane can thus be related to the substrate plane homography  $\mathbf{H}_s$  by

$$\mathbf{H}_b\mathbf{x}_b \cong (\mathbf{H}_s + \lambda\mathbf{M})\mathbf{x}_b \cong \mathbf{x}_b'$$

$$\mathbf{H}_s\mathbf{x}_b + \lambda\mathbf{M}\mathbf{x}_b \cong \mathbf{x}_b' \quad (10)$$

where  $\mathbf{x}_b = (x_{b1}, x_{b2}, x_{b3})^T$ ,  $\mathbf{x}_b' = (x'_{b1}, x'_{b2}, x'_{b3})^T$  are the pair of corresponding image positions projected from the bump plane, both expressed in homogeneous coordinates. In addition, the image correspondences over the substrate plane are related to  $\mathbf{H}_s$  by

$$\mathbf{H}_s\mathbf{x}_s \cong \mathbf{x}_s' \quad (11)$$

where  $\mathbf{x}_s = (x_{s1}, x_{s2}, x_{s3})^T$ ,  $\mathbf{x}_s' = (x'_{s1}, x'_{s2}, x'_{s3})^T$  are the pair of corresponding image positions projected from the substrate plane, again both expressed in homogeneous coordinates.

Equations (10) and (11) together allow the image point correspondences  $(\mathbf{x}_b, \mathbf{x}_b')$  from the bump plane  $\Pi_b$  and the image point correspondences  $(\mathbf{x}_s, \mathbf{x}_s')$  from the substrate plane  $\Pi_s$  to be used together for determining  $\mathbf{H}_s$ , and in turn  $\mathbf{H}_b$  through Equation (9).

More precisely, expressing  $\mathbf{H}_s$  as  $\mathbf{H}_s = [\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3]^T$ , for every pair  $(\mathbf{x}_s, \mathbf{x}_s')$  of correspondences from the substrate plane, we have a constraint for  $\mathbf{H}_s$  from Equation (11) which can be simplified to

$$\begin{bmatrix} \mathbf{0}^T & -x_{s3}'\mathbf{x}_s^T & x_{s2}'\mathbf{x}_s^T \\ x_{s3}'\mathbf{x}_s^T & \mathbf{0}^T & -x_{s1}'\mathbf{x}_s^T \end{bmatrix} \begin{bmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \\ \mathbf{h}_3 \end{bmatrix} = \mathbf{0}$$

or

$$\begin{bmatrix} \mathbf{0}^T & -x_{s3}'\mathbf{x}_s^T & x_{s2}'\mathbf{x}_s^T & 0 \\ x_{s3}'\mathbf{x}_s^T & \mathbf{0}^T & -x_{s1}'\mathbf{x}_s^T & 0 \end{bmatrix} \begin{bmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \\ \mathbf{h}_3 \\ \lambda \end{bmatrix} = \mathbf{0} \quad (12)$$

for some  $\lambda$ .

In addition, for every pair  $(\mathbf{x}_b, \mathbf{x}_b')$  of correspondences from the bump plane, we have another constraint for  $\mathbf{H}_s$  from Equation (11) which can be expressed as

$$\begin{bmatrix} \mathbf{0}^T & -x_{b3}'\mathbf{x}_b^T & x_{b2}'\mathbf{x}_b^T \\ x_{b3}'\mathbf{x}_b^T & \mathbf{0}^T & -x_{b1}'\mathbf{x}_b^T \end{bmatrix} \begin{bmatrix} \mathbf{h}_1 + \lambda\mathbf{m}_1 \\ \mathbf{h}_2 + \lambda\mathbf{m}_2 \\ \mathbf{h}_3 + \lambda\mathbf{m}_3 \end{bmatrix} = \mathbf{0}$$

where  $\mathbf{M} = [\mathbf{m}_1, \mathbf{m}_2, \mathbf{m}_3]^T$ . The above equation can be reduced to

$$\begin{bmatrix} \mathbf{0}^T & -x_{b3}' \mathbf{x}_b^T & x_{b2}' \mathbf{x}_b^T & p_1 \\ x_{b3}' \mathbf{x}_b^T & \mathbf{0}^T & -x_{b1}' \mathbf{x}_b^T & p_2 \end{bmatrix} \begin{bmatrix} \mathbf{h}_1 \\ \mathbf{h}_2 \\ \mathbf{h}_3 \\ \lambda \end{bmatrix} = \mathbf{0} \quad (13)$$

where

$$\begin{cases} p_1 = [\mathbf{0}^T & -x_{b3}' \mathbf{x}^T & x_{b2}' \mathbf{x}_b^T] \begin{pmatrix} \mathbf{m}_1 \\ \mathbf{m}_2 \\ \mathbf{m}_3 \end{pmatrix} \\ p_2 = [x_{b3}' \mathbf{x}^T & \mathbf{0}^T & -x_{b1}' \mathbf{x}_b^T] \begin{pmatrix} \mathbf{m}_1 \\ \mathbf{m}_2 \\ \mathbf{m}_3 \end{pmatrix} \end{cases}$$

Equations (12) and (13) allow all image correspondences be put together to estimate the parameters  $\{\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3, \lambda\}$ , i.e.,  $\{\mathbf{H}_s, \lambda\}$ , which in turn through Equation (9) allows  $\mathbf{H}_b$  to be estimated as well. With the two homographies  $\mathbf{H}_b$  and  $\mathbf{H}_s$  available, the overall bump height in terms of the norm of the paraplantar disparity matrix (PDM)  $\mathbf{D}$  can be determined from Equations (3) and (4). On the homograph estimation process, we adopt Direct Linear Transformation [6] as the kernel algorithm under a coarse-to-fine formulation to conduct the estimations.

In summary, while the biplanar-disparity method estimates  $\mathbf{H}_b$  and  $\mathbf{H}_s$  separately from either the image point correspondences from the bump plane or those from the substrate plane, the revised paraplantar-disparity method allows the two homography matrices be estimated simultaneously using both sets of correspondences. The revised method also enforces that the two homographies are induced from parallel planes, which is what the imaging configuration demands in practice. With more correspondences available for the estimation of either homography, and with the parallel-plane constraint enforced, the two homographies and in turn the overall bump height are estimated with higher accuracy despite disturbance from image resolution and other errors.

## 3 Experiments

### 3.1 Synthetic Data Experiments

Synthetic data experiments have the essence that ground truth about the true solution is available for assessing the quality of the experimentation. We have

carried out extensive synthetic data experiments to evaluate the performance of the Paraplantar Disparity Method, and the evaluation was done at both the level of estimating the homographies, since that is where the method has an impact, and the level of the final outcome, which is the norm of the disparity matrix (BDM or PDM). Monte Carlo simulation [12] was conducted to examine the variance of the estimates.

The parameters used for the settings of the two cameras, as tabulated in Table 1, were set as close to those in the real image data experiments as possible. The parameters include the intrinsic and extrinsic parameters of the cameras, as well as the alignment of the wafer bumps relative to the cameras.

Table 1: Imaging parameters used in synthetic data experiments.

Parameters of the synthetic system	
Bump diameter	120 micron
Bump-to-bump distance	70 micron
Number of bumps per inspection	20
Distance from wafer to camera frame	91mm
Angle between two light sources	15(degree)
Resolution of CCD	7.4 micron $\times$ 7.4 micron

With an imaging setting as stated in Table 1, we got 20 pairs of image point correspondences over the bump plane, denoted by  $\{(\mathbf{x}_b, \mathbf{x}_b')\}$ , as well as 20 pairs of the image point correspondences on the substrate plane, denoted by  $(\mathbf{x}_s, \mathbf{x}_s')$ . Zero-mean Gaussian noise of 1pixel standard deviation was added to the image positions of the correspondences, and the noisy data was used for homography estimation under the BDM-based and PDM-based methods respectively. A certain point on the center image (the image captured by the camera that faces top-down to the wafer) was projected to the lateral image (the image captured by the camera that faces the wafer laterally) by the bump-plane homographies  $\mathbf{H}_{b1}$  and  $\mathbf{H}_{b2}$  respectively estimated by the two methods. 100 trials were made. The noise-free output should be a point at the image position (16, 5) on the lateral image, but due to the added Gaussian noise, the actual output was a family of points around the position (16, 5). From Figure 4 it could be observed that the family of points induced by PDM-based method was less sparse than that of the BDM-based method, revealing that the PDM-based method is more robust in estimating the homographies.

Then we fixed the bump height at 60microns, and added the same Gaussian noise to the image positions of the correspondences. The norms of the BDM and PDM matrices were then computed respectively as reflections of the overall bump height under the two methods. With

noise-free data both norms appeared constant at all trials, as shown in Figure 5(a), but the Gaussian noise resulted in fluctuation of the actual output. Figure 5(b) and 5(c) show how the norms of the BDM and PDM matrices varied at different trials. It can be observed that the norm of the PDM was more robust against noise in the image point correspondences.

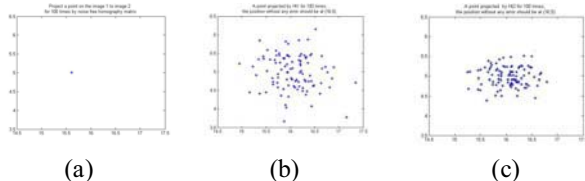


Figure 4: Comparison of the robustness of the BDM and PDM based methods in homography estimation. (a) The ideal image projection. (b) The distribution of image projections under the BDM based method. (c) The distribution of image projections under the PDM based method.

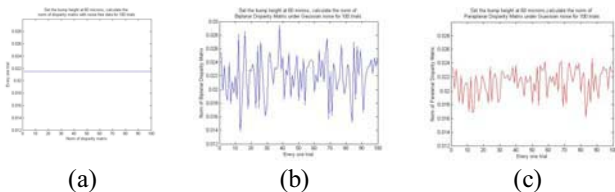


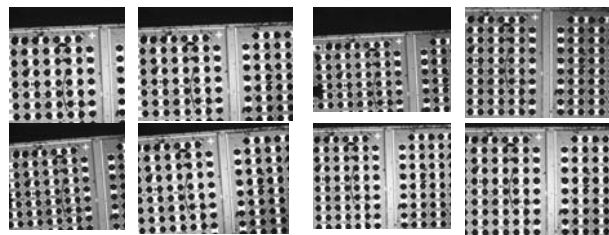
Figure 5: Comparison of the robustness of the BDM and PDM based methods on overall bump height determination. (a) The norm of both BDM and PDM under noise-free condition. (b) The norm of BDM under Gaussian noise. (c) The norm of PDM under Gaussian noise.

### 3.2 Real Image Experiments

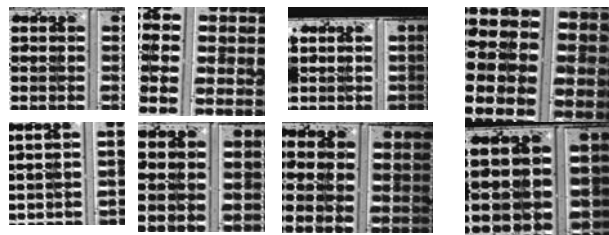
We have also conducted experiments on real wafers to compare the performance of the BDM and PDM based methods. Though unlike the synthetic data there was difficulty in obtaining ground truth about real data for evaluating estimation accuracy, the experiments could help examine estimation consistency (under different placement perturbations of the wafer) as well as whether the estimated values are close to the expectation or not.

We used the Sentec-405 CCD camera (752×582) and Navitar Zoom 6000 lens for image capture in the experiments. We took 8 datasets for the same wafer, each with the wafer’s position and orientation perturbed with a random translation in the range of 500 microns on the plane of the wafer, and a random rotation in the range of 2 degrees about the normal of the wafer plane, both about a particular central position and orientation of the wafer. The spatial perturbation was to simulate the non-repeatability of the wafer-feeding mechanism in the

manufacturing process. The image data are shown in Figure 6, and details of the processing steps are similar to that described in [4].



(a) 8 pictures captured by the center camera



(b) 8 pictures captured by the lateral camera

Figure 6: Image data used in real image experiments.

The norms of the BDM and PDM were both estimated from the data. Ideally, as the same wafer was imaged, all sets of image data should give the same overall bump height description and thus the same norms of the BDM and PDM. However, because of the geometric disturbances which resulted in change in the image data, the norms fluctuated across the various datasets. Table 2 listed out the norms under the BDM and PDM based methods, and Figure 7 has the norms plotted against the datasets. The norms had values as expected, showing that both methods gave reasonable results. It can also be observed that the variance of the PDM norms was reduced to one third of that of the BDM norms, which shows that the PDM method is generally more robust against placement change of the wafer than the original BDM method.

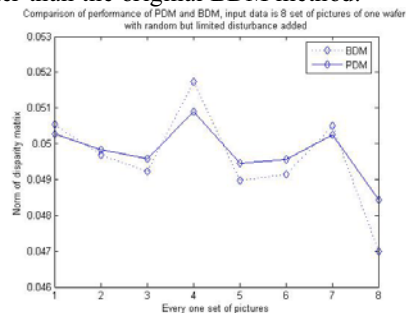


Figure 7: Norms of BDM and PDM under different placements of the wafer.

Table 2: Norms of the disparity matrices under the BDM and PDM based methods.

Dataset No.	1	2	3	4	5	6	7	8	Variance
BDM norm	0.0505	0.0497	0.0492	0.0517	0.0490	0.0491	0.0505	0.0470	19.384e-007
PDM norm	0.0503	0.0498	0.0496	0.0509	0.0495	0.0496	0.0503	0.0484	5.4857e-007

## 4 Conclusion and Future Work

This paper describes a new method for height inspection of a whole array of wafer bumps in parallel. It is based on a method we previously proposed, which was about an imaging configuration design that allowed the height information of multiple wafer bumps be picked up at the same time in image data despite the textureless and mirror-like nature of the wafer bumps. The height information was also made accessible, as the norm of a matrix, without going through explicit 3D determination.

However, the matrix was determined from two image-to-image mappings – the homography induced by the plane in 3D that contains most if not all of the bump peaks, and the homography induced by the bottoms of the bumps. Errors in estimating the two homographies due to limited number of image correspondences will propagate to the matrix and induce errors to the measurement of the bump heights.

This paper describes a method that allows the two separate sets of image correspondences – each for estimating one homography – be combined together for estimating both homographies simultaneously. The method is made possible by exploiting a constraint that holds in practical imaging condition, which is that the two planes that induce the two homographies are parallel and close to each other in space in comparison with the imaging distances. Experiments have shown that this paraplanner constraint allows the homographies and in turn the bump height information be estimated with substantially higher accuracy, for a range of bump heights used in practice, despite perturbations from limited image resolution and wafer placement nonrepeatability.

The inspection system consists of only two sets of cameras and light sources. No moving part is involved in the entire inspection process. In other words, nonrepeatability of positioning devices and limitation of the inspection speed from mechanical scanning are not in the picture. With such simplicity and features, the system is well suited for inspecting wafer bumps, especially in cases where only the height of the wafer bumps not the whole shape profile of the bumps is required. Massive trials in manufacturing environment will be the future work.

## Acknowledgement

The work described in this paper was substantially supported by a grant from the Innovation and Technology Commission of Hong Kong Administrative Region, China, under an Innovation and Technology Fund with Project Code UIM/111.

## References

- [1] Rideout, E., “BGA inspection”, Proceedings on Electronics Manufacturing Technology Symposium, Vol.1, pp. 388.
- [2] Burton, D.R., Lalor, M.J. and Atkinson, J.T., “The growth of modern interferometry for industrial inspection”, IEEE Colloquium on Active and Passive Techniques for 3-D Vision, pp. 31-34, Feb. 1991.
- [3] Sakuma, H., “Interferometry for rough surface”, International Conference on Lasers and Electro-Optics Europe, pp. 230-235, Sept. 2000.
- [4] Dong Mei, Ronald Chung, “Height Inspection of Wafer Bumps without Explicit 3D Reconstruction” IS&T/ SPIE 18th Annual Symposium on Electrical Imaging at San Jose, USA. in January, 2006.
- [5] Zhou Chuan, Tan Dalong, Zhu Feng “A planar homography estimation method for camera calibration”, Computational Intelligence in Robotics and Automation, 2003. Proceedings. 2003 IEEE International Symposium
- [6] Richard Hartley and Andrew Zisserman, “Multiple view Geometry”, Cambridge University Press, second edition, p138-150
- [7] M.A. Fischler and R.C. Bolles, “Random sample consensus: A paradigm for model fitting with application to image analysis and automated cartography”, Communications of the ACM, Vol. 24(6), pp. 381-395, 1981.
- [8] Onn, R., Steinhardt, A.O. and Bojanczyk, A.W., “The hyperbolic singular value decomposition and applications”, IEEE Transactions on Signal Processing, Vol.39, pp.1575 - 1588, July 1991.
- [9] Harrison M. Wadsworth, “Statistical methods for engineers and scientists”, McGraw-Hill Publishing Compan, pp16.2-16.22,1998.
- [10] Luong, Q.-T. and Faugeras, O.D., “Determining the fundamental matrix with planes: instability and new algorithms”, IEEE Proceedings on CVPR'93., 15-17, pp.489 - 494, June 1993.
- [11] Etienne Vincent and Robert Laganeire, “Detecting planar Homographies in an image pair”, Proceeding on 2nd International Symposium on Image and Signal Processing and Analysis, pp. 182-187, Pula, Croatia, June 2001.
- [12] Greg Kochanski, “Monte Carlo Simulation”, <http://kochanski.org/gpk/>
- [13] Golub, G. H., and Van Loan, C.F. “Matrix Computations”, Johns Hopkins University Press, 2nd ed., 1989.
- [14] C. Harris and M. Stephens, “A combined corner and edge detector,” in Proc. 4th Alvey Vision Conference, Manchester, 1988, pp. 147–151.