3D INSPECTION OF WAFER BUMP QUALITY WITHOUT EXPLICIT 3D RECONSTRUCTION

ZHAO YANG

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF PHILOSOPHY

IN

AUTOMATION AND COMPUTER-AIDED ENGINEERING

©THE CHINESE UNIVERSITY OF HONG KONG
JULY 2004

The Chinese University of Hong Kong holds the copyright of this thesis. Any person(s) intending to use a part or whole of the materials in the thesis in a proposed publication must seek copyright release from the Dean of the Graduate School.



A Theres Scientific in Page of Therese and the Commission of the C

office and the amount of

Астомучной спр Сприменный Каримент

de l'un Chaixean University de l'Ionne 1000

The Circumstation of the Committee of the Committee of the charter of the Committee of the

proposed pulsarious areas and the contract of the contract of

Acknowledgements

I would like to thank my supervisor, Prof. Ronald CHUNG Chi kit and Prof. XU Yangsheng for the interesting problem they introduced to me and for providing useful technical discussions over the course of my research. I would also like to express my sincere appreciation to Prof. LIU Yunhui, Prof. LI Youfu of the City University of Hong Kong and Prof. Edmund Lam of the University of Hong Kong for their helpful suggestions and support.

I would like to thank ASM automation assembly Ltd. for their project and their great help in the experiments and their support for my research. I would like to say thank you to Mr. WH Leung, Dr. Kenneth Fung, Ms. Brenda Lau, Dr. WANG fan for their kind help and advices in every student meeting.

I would like to remember all the colleagues and friends who shared with me a lot of happiness during my study. To name only some of them: Mr. WANG Wei, Mr. CHENG Jun, Mr. HE Yong, Ms. YUAN Ding, Mr. Ben Chun, Mr. LIANG Bodong, Mr. YUAN Meng, Dr. OU Yongsheng, Dr. SHI Xi, Dr. K.K. Lee, Mr. Kyle, Mr. Tony, Miss LI Yu jia, Mr. CHEN Tai, Mr. CHEN Zhiyong, Ms. QU Cong, Mr. LIU Tong, Mr. YU Mao lin, Mr. HU Guoqiang, Mr. DAI Ruoli, Mr. WANG Hesheng, Miss LIU Lu, Mr. WEI Peng, Miss LI Dandan.

I wish to express my deepest gratitude to Mr. WANG Xin and my family for their training of my discipline and for their continuous encouragement.

Abstract of thesis entitled:

3D Inspection of Wafer Bump Quality without Explicit 3D Reconstruction

Submitted by ZHAO Yang

for the degree of Master of Philosophy
in Automation and Computer-aided Engineering
at The Chinese University of Hong Kong in
July 2004.

One goal of the semiconductor industry is to achieve smaller-size chips and circuitry. One way of achieving that is to, instead of achieving connections on PCB [19](Printed Circuit Board) between packaged chips, allow die-to-die or wafer-to-wafer direct bonding. The advanced packaging technology named wafer bumping or wafer-level packaging (WLP)[20] is exactly that: it could reduce device size, improve their performance and manufacturability, and ultimately reduce cost.

Alongside the development of the WLP technology is a demand for improved process control as well as quality assurance. In particular, wafer-to-wafer bonding requires solder bumps to be first established at corresponding positions of both wafers, just like in the case of PCBs. However, because of the much reduced size of the circuitry in wafers or dies in comparison with that of PCBs, the wafer bumps are of much smaller size, typically with a diameter of only tens of microns. Bumps of such size are difficult to have their 3D shape reconstructed with enough accuracy, yet even tiny errors on their heights and so on will be a hazard to the bonding process. Because of the minute size of the bumps, visual inspection technologies widely used for PCB bumps are not applicable. 3D inspection of wafer bumps is a bottleneck to the WLP technology.

One particular and important inspection needed in WLP is to check whether all the bumps on a wafer are of the same and specified height, so as to prevent too strong or too loose contact of a few bumps in the bonding process. There have been a few methods developed for it, but they all require explicit 3D reconstruction which is an expensive process especially for dimensions of such a miniature scale.

In this thesis, we introduce a novel methodology that could inspect the heights of the wafer bumps without conducting explicit 3D reconstruction. The methodology incorporates the reflective nature of solder bumps, a transformed version of the stereo vision concept, appropriate illumination and positioning of cameras, the SVD [31](Singular Value Decomposition) tool, and a 3 by 3 measurement matrix we call Biplanar Disparity Matrix that could summarize the global information about the heights of the wafer bumps. Experimental results with synthetic and real data show that the methodology could give promising results.

摘要:

利用非三位重建方法对晶片上触点的三维检测 哲学硕士学位申请人:赵阳 香港中文大学 自动化与计算机辅助工程 二零零四年七月

在半导体工业中,我们不断追求尽可能的生产出尺寸小的芯片与电路板。除了使用印刷电路板外,我们还有一个途径就是利用模板与模板、芯片与芯片的直接连接。 一种现代化的集成技术,被称为晶片触点连接,或者叫做晶片级集成就是如此。这种继承方法可以缩小产品的尺寸,提高产品的工作效果,最终节省投资。

随着优化过程控制技术以及产品质量可靠性的发展与提高,晶片级集成技术的开发势在必行。尤其是,晶片与晶片的可靠连接,需要以晶片上突起触点的正确的三维位置为基础,正如印刷电路板的原理一样。然而,由于晶片为了适应印刷电路板尺寸的日益减小而缩小尺寸,实际上,通用的晶片上的突起的直径只有几十个微米。由于被检测的尺寸如此之小,传统上用来检测印刷电路板的三维重建的方法是很难用在晶片触点的检测上的。因此,晶片凸起的三维检测造成了阻碍晶片级集成工业发展的瓶颈。

在晶片级集成电路生产过程中,一项最重要的指标就是提前检测晶片上凸起触点的 高度是不是标准的高度,这样可以阻止在集成过程中,高度过低或者过高的触点 影响整个电路的连接。事实上,目前已经有一些检测的方法被工业界应用,但是这 些方法都需要完整的三维重建,这个过程既繁杂,又耗费时间与金钱。 在这篇论文里,我们提出了一种新的检测方法,这种发放不需要完整的三维重建也能过实现对微米级晶片凸起的检测。 我们的方法理论结合了晶片凸起的反射性质,也同时利用了立体视觉概念的变形,以及合适的照明装置,和恰当的相机的位置,我们还利用了奇异阵分解等工具,综合提出了一个3乘3的矩阵,叫做双平面间差异矩阵,作为新的度量触点高度的标准。实验结果与仿真试验都证明,我们的方法有着很大的开发价值,和美好的前景。

Contents

Contents

1	Int	roduct	ion	1						
	1.1	Objec	tives of the Thesis	1						
	1.2	Wafer	bumping inspection by using Biplanar Disparity approach	2						
	1.3	Thesis	S Outline	4						
2	Background									
	2.1	What	is wafer bump?	5						
		2.1.1	Common defects of wafer bump	6						
		2.1.2	Literature review on exist wafer bump inspection method	11						
3	Mo	del 1:	the one camera model-Homography approach	21						
	3.1	The in	ntroduction of the theoretical base of model 1	21						
		3.1.1	The objective of model $1 \ldots \ldots \ldots \ldots$	21						
		3.1.2	Desires	22						
		3.1.3	Some background knowledge on Homography	22						
	3.2	Model	1- "Pseudo Homography" Approach	24						
		3.2.1	The description of the configuration of model 1 \dots	24						
		3.2.2	The condition of pseudo Homography	25						
		3.2.3	The formation of pseudo Homgraphy H	26						
	3.3	Metho	odology of treatment of the answer set	32						
		3.3.1	Singular Value Decomposition-SVD	32						
		3.3.2	The Robust Estimation	33						
		3.3.3	Some experimental results by using manmade Ping Pang							
			balls to test SVD[31] and Robust Estimation[24]	35						

		3.3.4	the measurement of the Homography matrix answer set .	37
	3.4	Prelin	minary experiment about model 1	43
	3.5	Probl	ems unsolved	47
4	Mo	del 2:	the two camera model-Biplanar Disparity approach	48
	4.1	Theor	retical Background	48
		4.1.1	the linearization of Homography matrix changes	49
		4.1.2	Problem Nature	51
		4.1.3	Imaging system setup	52
		4.1.4	Camera Calibration[13]	52
	4.2	Metho	odology	54
		4.2.1	Invariance measure	54
		4.2.2	The Geometric meaning of the Biplanar Disparity matrix	58
	4.3	RANS	SAC-Random Sample Consensus	64
		4.3.1	finding Homography matrix by using RANSAC[72][35] $$.	64
		4.3.2	finding Fundamental matrix by using RANSAC[73][34] $$.	65
	4.4	Harris	S Corner detection	65
5	Sin	nulatio	n and experimental results	67
	5.1	Simula	ation experiments	67
		5.1.1	Preliminary experiments	67
		5.1.2	Specification for the synthetic data system	71
		5.1.3	Allowed error in the experiment	71
	5.2	Real i	mages experiments	72
		5.2.1	Experiment instrument	72
		5.2.2	The Inspection Procedure	74
		5.2.3	Images grabbed under above system	75
		5.2.4	Experimental Results	81
6	CO	NCLU	SION AND FUTURE WORKS	83
	6.1	Summ	ary on the contribution of my work	83
	6.2	Some	Weakness of The Method	84
	6.3	Futur	e Works and Further Development	8/

6.3.1	About the synthetic experiment .							84
6.3.2	About the real image experiment							85
Bibliography								87

List of Tables

3.1	Intrinsic parameter of the camera	45
5.1	Intrinsic parameter of the camera	71
5.2	Extrinsic Parameter of the cameras and CCD resolution $\ \ldots \ \ldots$	73
5.3	different norms with respect to different wafer	81
5.4	changing ratio of the norms' value within the two groups \dots	82
	A significance is occurred to a sugarity objects for on-	

List of Figures

1.1	imaging set up for Biplanar Disparity Measurement	3
2.1	procedure of making wafer bumps	6
2.2	BGA bumps	7
2.3	the 3D image of a real bump	7
2.4	the configuration of a wafer-level packaging	8
2.5	bump malformation	9
2.6	bump bridging	9
2.7	bump misplacement	10
2.8	the 3D defects in solder bumps	10
2.9	the principle of laser scan system	13
2.10	the configuration of moire interferometry inspection system	14
2.11	3D Z-topography of bare die showing a bump defect	15
2.12	Volume pixel acquisition (VPA) methodology captures X, Y and	
	Z data for each pixel	16
2.13	Depth Recovery Triangulation	18
2.14	Principle of confocal system	19
	confocal measurement used by August Ltd	20
3.1	the configuration of model 1	24
3.2	a single camera observing a moving object is equivalent to a	
	"counter-moving" camera observing a stationary object	27
3.3	detail of the "counter-moving" camera observing a stationary object	28
3.4	the formation of pseudo Homography	29
3.5	the formation of pseudo Homography	30

3.6	the formation of pseudo Homography	31
3.7	the formation of pseudo Homography	32
3.8	A distribution in two dimensions fitted to a straight line \dots .	33
3.9	The reference image and the specimen image with an outlier $$. $$	36
3.10	Pointing out the outlier by using robust estimation	37
3.11	the separation of subspace	39
3.12	Flow chart of inspection procedure by model $1 \ldots \ldots \ldots$	44
4.1	theoretical analysis on the change in norm of Biplanar Disparity	
	with respect to bump height change from 20 to 100 microns[29]	62
5.1	Norm of D in the absence of any error or uncertainty $\ \ldots \ \ldots$	68
5.2	Norm of D in the presence of image resolution error \dots	68
5.3	Norm of D in the presence of image resolution error and bump	
	position variation uncertainty	69
5.4	Norm of D in the presence of image resolution error, brightest	
	point determination uncertainty, and bump position variation	
	uncertainty	69
5.5	Norm of D in the presence of image resolution error, bump po-	
	sition variation uncertainty and wafer-plane transformation and	
	wafer-tilting transformation	70
5.6	Norm of D in the presence of image resolution error, brightest	
	point determination uncertainty, bump position variation uncer-	
	tainty, and wafer-tilting transformation.	70
5.7	for center imaging part	73
5.8	for side imaging part	74
5.9	Real image group 1-center images:1,2	75
5.10	Real image group 1-center images:3,4	76
5.11	Real image group 1-center images:5,6	76
5.12	Real image group 1-side images:1,2	77
5.13	Real image group 1-side images:3,4	77
5.14	Real image group 1-side images: 5,6	78

5.15	Real image group 2-center images:1,2	78
5.16	Real image group 2-center images:3,4	79
5.17	Real image group 2-center images:5,6	79
5.18	Real image group 2-side images:1,2 \dots	80
5.19	Real image group 2-side images:3,4	80
5.20	Real image group 2-side images:5,6	81
5.21	two groups of D. Norm value about two pieces of Die with differ-	
	ent bump size	82
6.1	Norm of D in the presence of image resolution error, brightest	
	point determination uncertainty, bump position variation uncer-	
	tainty, and wafer-tilting transformation, under a higher image	
	resolution 1000 X 1000	85

Chapter 1

Introduction

The goal of modern semiconductor design is to achieve shorter electron pathways for increased speed, power and total device functionality. Advanced packaging technology, such as wafer bumping and wafer-level packaging, is a critical area in semiconductor manufacturing. It can improve device performance and manufacturability, and ultimately reduce cost. Ongoing developments in wafer scale technology continue to address these goals while maintaining or decreasing device size. During the process of packaging, to get the wafer-wafer connection effective and efficient, all the bumps on the wafer are required to have the same specific height. Whose height is too low or too high will damage the whole board into invalid, and should be examined out.

1.1 Objectives of the Thesis

Traditional methods, such as laser scan, confocal, moire interferometery, and traditional stereo vision, all of which require the use of moving parts, stringent light projection system, and exact mirror and lenses for scanning and focusing which are all costly. However, in many applications, detailed 3D shape profile of the target surfaces which is the bump surfaces in the case of wafer inspection is not needed. What is needed is to measure whether the bump heights on the whole meet the specifications or not, as well as to identify the individual bumps

that are too tall or too short. This work aims at coming up with a system that does just that, requires an integrated study of an innovative approach in the lighting, optics and image processing that is suitable for development in the academic environment without going through explicit 3D reconstruction and thus save operation speed and hardware cost of the whole system.

1.2 Wafer bumping inspection by using Biplanar Disparity approach

In this thesis, we propose a new approach called Biplanar Disparity measurement. This approach mainly concerns on measuring the change of norm of the Biplanar Disparity matrix, thus measuring the change of bump's height. Because the bumps' height is the only criterion of examining whether some bumps are golden or defects. The imaging system is shown in the below Fig.1.1:

Red camera and red parallel lighting are all perpendicular to the wafer working plane, while blue camera and blue parallel lighting are symmetrically placed with the red lighting angle bisector. Then, in the image from blue camera, we only concern on the blue point caused by the blue lighting, similarly, in the image from the red camera, we only concern on the red point caused by the red lighting. In the two correspondening[15] images, the red point and blue point are the same point in 3D-the peak point of the bump. So, they can form a pair of correspondence points. And the Homography matrix computed from these points is about every peak plane formed by the peak points of the bumps on the specimen wafer.

With some constraints, we can respectively find some other points in the two images correspondingly and in 3D on the substrate, which can be computed to get the Homography matrix about the substrate. Using these two Homographies, we induce a new variant- the Biplanar Disparity matrix D.

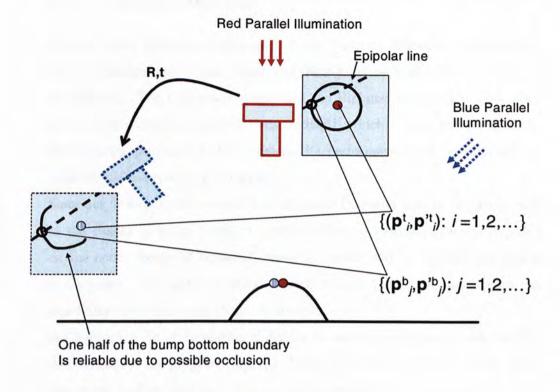


Figure 1.1: imaging set up for Biplanar Disparity Measurement

D's expression is:
$$\mathbf{D} = \frac{\hat{\mathbf{H}}_t}{\frac{w_t}{\alpha}} - \frac{\hat{\mathbf{H}}_b}{\frac{w_b}{\alpha}} \tag{1.1}$$

Obviously from the expression, D is related only to both of the Homographies about the bumps' peak plane and the bottom plane. So, the Biplanar Disparity matrix represents the difference between the two planes which indicates the bumps' height. The assumption for this approach is that most of the bumps on the specimen wafer are coplanar, those bumps which have quite different height will be got rid of before the Homography-estimation[16] loop knot. And this is also true in the actual wafer bump inspection industry.

1.3 Thesis Outline

Firstly, before detailed examination of the Biplanar Disparity Measurement method, the concept of wafer bump and some previous work will be introduced in Chapter2. The final model presenting the Biplanar Disparity methodology is based on an initiative one-camera model[61] which is also derived from the Homography principle of stereo vision. Its performance will be evaluated by some simulation results in Chapter3.

Since our focus is in the sensitivity of Biplanar Disparity matrix D with respect to the change of bump height, the effect of height change of bumps imposed on the norm change of Biplanar Disparity matrix will be studied thoroughly in Chapter 4. Extensive simulations have been conducted to investigate the sensitivity of the Biplanar Disparity matrix.

In Chapter5, a set of simulation experiments has been designed to validate the effectiveness of the proposed Biplanar Disparity measurement approach. Both theoretical analysis and real data experiment results will be shown. It will be found that the Biplanar Disparity matrix D works very well for detecting the change in wafer bump height.

Finally, conclusions will be made and future research directions are suggested.

Chapter 2

Background

2.1 What is wafer bump?

The goal of modern semiconductor design is to achieve shorter electron pathways for increased speed, power and total device functionality. Advanced packaging technology, such as wafer bumping and wafer-level packaging, is a critical area in semiconductor manufacturing. It can improve device performance and manufacturability, and ultimately reduce cost. Ongoing developments in wafer scale technology continue to address these goals while maintaining or decreasing device size. The Fig.2.1 can simply describe nowadays the product process of wafer bumping technologies.

After this industrial process, the bumps on the die are almost finished. In the real world inspection process, the bumps on the wafer needed to be inspected are mostly like this kind shown below, called ball grid array (BGA).

Fig.2.2 shows the stereo image of BGA bumps [42]under microscope. Generally speaking, the top of the bumps is the connection in the wafer level packaging industry.

Despite recent industry downturn, the market for wafer-level packaging is predicted to rise significantly in the coming few years. Alongside the development of this technology is a demand for improved process control and quality assurance. Automated machine vision inspection is frequently needed for such purpose.

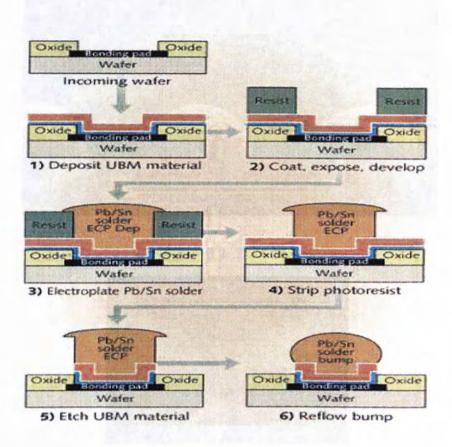


Figure 2.1: procedure of making wafer bumps

2.1.1 Common defects of wafer bump

Typically, 2D inspection is needed for detection, review and classification of defects. There are also sustained advancements in the manufacturing process that demands better vision inspection algorithms to maintain high throughput

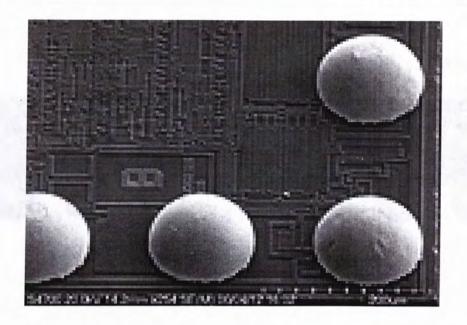


Figure 2.2: BGA bumps

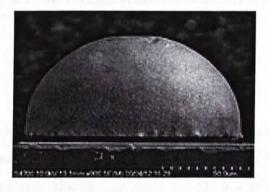


Figure 2.3: the 3D image of a real bump

and accuracy. For wafer-level packaging, 3D metrology and inspection is also required. Critical measurements are height, diameter (for special bumps), length and width (for rectangular bumps), coplanarity and true position. Of all of these measurements, by far height and coplanarity are two key features. Bumped die whose I/O contact points fail to conform to strict height and coplanarity dimensional tolerance represent a potential electrical failure downstream. A profound bump height or coplanarity problem can cause a probe needle to break during the probing process resulting in a massive financial lost.

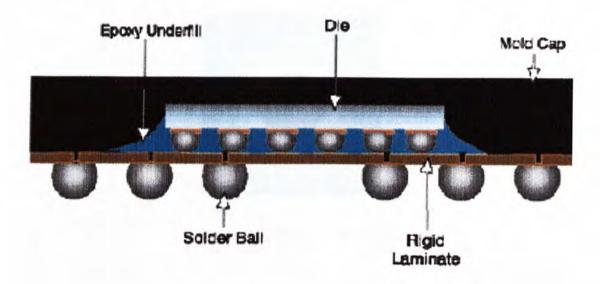


Figure 2.4: the configuration of a wafer-level packaging

2D defects

Some defects, such as missing bumps or bridges between bumps, which are shown below, immediately affect the functionality of the final package. Other defects, such as bump malformation, bump misplacement and inter-bump surface contamination and change in diameter of bumps will affect the product's long-term reliability. Typically, a CCD camera is used to collect images for these measurements and defect identification. The combination of dark field and bright field lighting is used to provide the correct illumination to high-light the area or the feature of interest. Inspection methodologies covering 2D features are fairly standard whether they are for metrology or surface defects. Sometimes, people use image processing to extract the features, and with some signal processing traditional method to deal with the images, thus find these defects. These methods do not require the bumps' three dimensional information, such as height, so we call these kinds of defects- 2D defects.

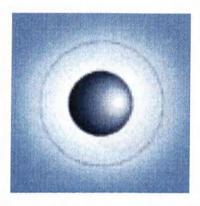


Figure 2.5: bump malformation

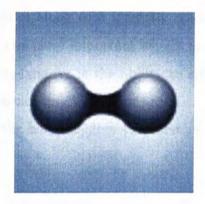


Figure 2.6: bump bridging

3D defects

From Fig.2.1 we know that, during the process of wafer-level packaging production, the main part is to realize the bump-bump connection. If the height of the bump can not meet the requirement, simply understood as not as the same height as the golden bump, that will lead to the whole wafer to be invalid. Thus, the inspection of height of bumps and coplanarity is a must before packaging. In this stage, the general 2D measurement can not get to the destination. We need

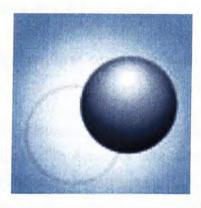


Figure 2.7: bump misplacement

to apply 3D measurement. However, 3D measurement methodology is somewhat less standard than 2D across the industry, as there are several techniques that are used to provide 3D measurements. Lasers, standard 2D optics with auto focus, and 2D optics along with intensity readings (confocal) are among the techniques used. The most important aspect in 100percent 3D inspection of bump height and coplanarity is the ability to simultaneously achieve high throughput and high accuracy.

We need to measure the bump heights and examine if defects may occur with

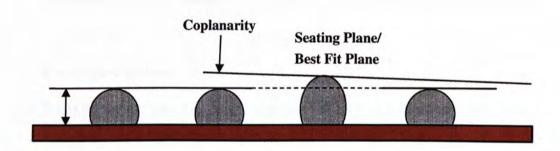


Figure 2.8: the 3D defects in solder bumps

bumps that are too high or too low. This requires and integrated study of an innovative approach in the lighting, optics and image processing that is suitable

for development in the academic environment.

2.1.2 Literature review on exist wafer bump inspection method

There are a number of companies developing 2D and 3D inspection machines for wafer level packaging. The dominant ones include August Technology in Minnesota, U.S.A., RVSI[50] in Massachusetts, U.S.A., and ICOS Vision Systems in Belgium. Each has introduced products in the wafer bump inspection market.

Because of the rapid evolvement in the semiconductor manufacturing process, continuous R and D is necessary to improve the visual inspection mechanism, predominantly with speed and accuracy. Generally speaking, there are 4 methods that have been used in Wafer Level Packaging industry: Laser Scan, Fast Moire Interferometry (FMI)[36], traditional stereo vision, and Confocal Microscopy. All four methods require the use of moving parts, stringent light projection system, and exact mirror and lenses for scanning and focusing which are all costly.

August Technology, for example, uses a confocal sensor in their 3D system that compromises on speed in exchange for better accuracy. ICOS uses stereo imaging that does not yet give good 3D profile information. RVSI has better 3D measurement techniques but not as good defect detection and classification.

Laser scan system

This technique relies on sensors to measure light reflected off a target onto either a pixelized array detector or a position-sensing detector[37]. The laser diode for the sensor projects a beam of light onto the target. The light-sensitive detector built into the sensor receives some of the reflected light and records the position of the reflected beam, along with a measurement of height. As the target (or the sensor) moves to another image point on the component, the position of the reflected light beam changes, and new data (height and position) are recorded

as a delta from the previous position.

The technology is somewhat limited for in-tray inspection, because the components need to be perfectly aligned for the laser beam to properly inspect them. The laser also scans only the areas of the component directly in-line with the path. Those areas between the laser path assume a comparable surface, which can lead to missed defects. Then too, laser triangulation measures only height (Z axis), and must be combined with an X-Y mechanism to obtain 3D or 2D data. As the target or sensor moves, the spinning mirrors and precision drives represent a complex system that can be adversely affected by mechanical wear and other problems.

Finally, the most limiting aspect of laser technology (in addition to speed) is undoubtedly spot size, the diameter of the "spot" made by the laser at the point of contact. If the diameter, for instance, is 30 microns, then the system cannot detect lateral features less than 30 microns. Such a limitation is acceptable for BGAs, but not so for smaller components, such as micro BGAs[52] and flip chips. Lasers[48] are also subject to the "speckle" effect (noise or interference in the image produced by scattering of the reflected beam).

The principle of laser scan[43][46] can be seen in figure Fig.2.9. The laser offered in this method is like a plane scanning from one side of the bump to the other. During this process, there is camera placed perpendicular to the substrate of the wafer grabbing several images. The more the pictures the camera grabbed, the more information we will get from the images. With these images, the bumps' 3-D information can be got.

Moire interferometry

This technology has been employed for many years in 3D profiling of difficult-to-measure parts in various industries other than electronics (e.g., automotive, aeronautics, and manufacturing equipment)[51]. The procedure entails: Illuminating a bump with a moire light pattern created by projecting a laser or

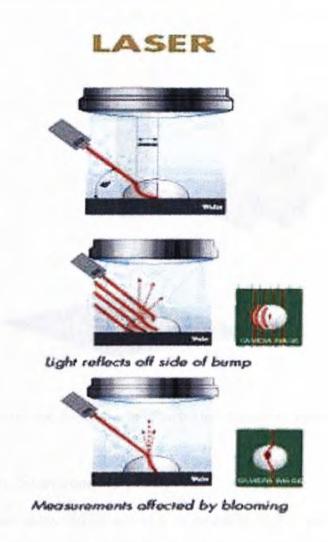


Figure 2.9: the principle of laser scan system

white light beam through a grid and capturing the resulting image with a CCD (charge coupled device) camera. Moving the light pattern a few microns and capturing another image. Repeat the process for a third time. In each case, light intensity and position readings are measured and processed by software to achieve a profile of the target object. In recent years, a few companies have tried to develop a Moire Interferometry inspection [50]tool for the electronics manufacturing environment, but have encountered several challenges. One company uses a scanning process, which requires that the component be conveyed in a tray or on tape beneath the camera and laser system, thereby limiting inspection throughput. Scanning speed is said to be 125mm per second.



Figure 2.10: the configuration of moire interferometry inspection system

Fast Moire Interferometry (FMI)

FMI [40](fast moire interferometry) is an evolution of the basic moire technology; and as such, it raises moire interferometry to a level where the process is clearly superior to other techniques for inspecting bumped packages. The result is three times the inspection speed of existing moire interferometry systems. Fig.2.10 depicts the configuration for FMI.

The features and performance characteristics of FMI inspection[51] technology can be summarized as follows:

FMI uses a single CCD camera, combined with a projector that diffuses a white light pattern on the component or tray of components being inspected through a grating mounted in the projection assembly. The light source provides an intense beam using an optical fiber coupled with an aspherical lens located between the light source and the grating. A grating of light is thus projected

onto the component, following the Z-topography of the surface.

The color CCD camera is situated directly above the component at a 90 degree angle to the platform, while the projector is fixed at a 30 degree angle. (The camera and the projector are mounted on the same gantry system.) The high-resolution camera captures the image of the grating projection on the component, and converts the pixel data into digital output.

The system captures an entire field of view; therefore, no scanning is required. In fact, the field of view can be scaled up to encompass multiple components simultaneously. Data is thus acquired faster than with any other system on the market.

Unique to FMI is the fact that the grating, which creates the moire pattern on the target, is moved sequentially four times with high precision. As the grating moves, four different images are recorded, along with multiple levels of intensity. Algorithm software then converts the data into 3D and 2D images simultaneously.

FMI incorporates a proprietary methodology-volume pixel acquisition



Figure 2.11: 3D Z-topography of bare die showing a bump defect

(VPA)[44]-which captures X and Y, as well as Z data for each pixel Fig.2.12. Superior X-Y platform control is achieved through the use servo motors and encoders. The 2D x-y data depicts such conditions as the presence/absence of bumps, proper fiducial alignment, correct component markings and so on. The

captured X, Y, and Z data results in a "pass" or "fail" indication by the system FMI systems can inspect individual components or multiple components in a tray. When required, a magnetic transfer table is used to move the packages from the tray to the work area.

Using VPA, the volume of solder bumps and balls can be determined. FMI technology can also be employed to inspect the quality of the mold encapsulating both leaded and area array devices and packages.

While area array components seldom require interconnect diameters under 10 mils (0.010 in.), FMI systems can inspect bumps and balls as small as 40 microns in height.

Fast moire inspection equipment is available today as fully automated,

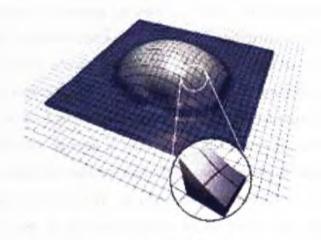


Figure 2.12: Volume pixel acquisition (VPA) methodology captures X, Y and Z data for each pixel.

programmable in-line systems for 3D and 2D inspection and as semi-automated batch systems for off-line process verification and product development. The equipment can also be designed as a module to be incorporated within other production equipment, such as a system that places solder balls on BGAs (ball grid arrays) or flip chips. The module would enable automatic inspection of the existence and linear placement of the balls, as well as ball height, volume, and substrate warpage.

In addition to packaging, FMI technology offers particular advantages in the 3D inspection of various electronic components, given the superior processing speed and the sub-micron precision achieved. Other applications include inspection of solder paste after stencil printing of circuit boards and inspection of wafer-level packages[17], namely solder bumps on 200mm and 300mm wafers. In the case of both solder paste and wafer bumps, FMI enables in-line line inspection without degrading throughput, though off-line batch inspection would be possible as well, depending on the line set-up and requirements of the manufacturer.

Stereo vision

Stereo vision [1],[2] employs two cameras, each positioned with a different line of sight to the object being imaged (a solder ball for instance). The effect is somewhat similar to human binocular vision - in other words, the eyes of Human Beings - in viewing the interconnection. The intersecting lines of sight to any point on the object are "matched" [5] and measured on a pixel-by-pixel basis to construct a map. Algorithms are then created to produce a 3D image that can be viewed from virtual or imaginary camera angles.

Stereo vision has significant limitations, one being that the calibration required between the two cameras is tedious. While stereo inspection systems can adjust for component size, the technology cannot easily accommodate free-form surfaces (flat surfaces with no edges for reference), and the shape of the component must be determined prior to the inspection process. Also, the requirement that one of the two cameras be fixed at a specific angle restricts the inspection region and presents difficulty in inspecting the entire component. In effect, when inspecting a solder ball, stereo vision results in only a single data point. So, if we want to know the 3-D profile of a wafer/Die, we need to reconstruct the three dimensional configuration of every single bump point by point explicitly. The principle can refer to the Fig.2.13 below.

If there are enough correspondence points (at least four pairs) on image I and

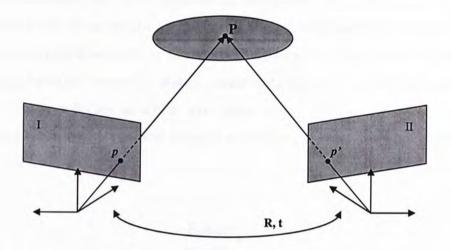


Figure 2.13: Depth Recovery Triangulation

image II, we can get the 3-D information of point P in the world coordinate plane.

Confocal

The process is based on producing a linearly polarized beam using a laser, the beam being transmitted through various optical elements mirrors, lenses, beam splitter, in hole plate, etc. - where it impinges a target, such as a solder bump, at a specific and minute location. (The beam at the impingement point on the target is about a micron in diameter.) The beam is then reflected back, filtered to produce a high signal-to-noise ratio, and passed through the beam splitter[44], where it is deflected at right angles to the path of the transmitted beam and is directed by a turning mirror to a photo detector. The photo detector converts the amount of light into digital output to a computer. The image composed from the various impingement points on the target is subsequently displayed on the computer monitor.

The method measures the amount of light reflecting directly off the top of any feature on a wafer surface. It correlates the light intensity with the vertical position of the sensor through a relationship defined by the confocal optics. Actual feature height is calculated by comparing the relative height of the feature to a reference surface selected by the user. The confocal method uses a continuous scanning motion to capture feature height data and its speed is independent of the number of bumps on die or on a wafer.

The problem with confocal imaging equipment, besides the high capital ex-

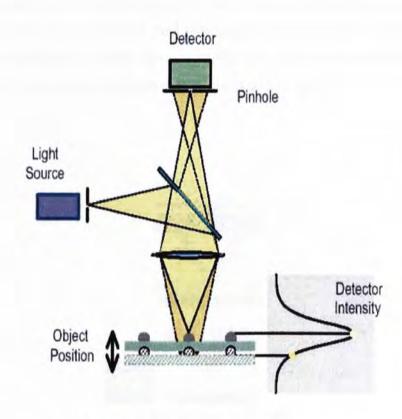


Figure 2.14: Principle of confocal system

pense, is the extreme degree of alignment that must be maintained in terms of the optics (mirrors, lenses, etc.). Vibration and shock, even through normal conveyor movement, can disturb the alignment and provide erroneous data. Setup time can also be lengthy. In addition, throughput is a major issue. Since confocal [23], [30], imaging is conducted at several Z-locations (at least 5 to 7

layers), inspection speed is greatly reduced. As a result, a sampling alternative is suggested instead of 100 percent inspection of the entire part.

In a word, all four methods require the use of moving parts, stringent light projection system, and exact mirror and lenses for scanning and focusing which are all costly. However, in many applications, detailed 3D shape profile of the target surfaces (the bump surfaces in the case of wafer inspection) is not needed. What is needed is to measure whether the bump heights on the whole meet the specifications or not, as well as to identify the individual bumps that are too tall or too short. This work aims at coming up with a system that does just that, requires an integrated study of an innovative approach in the lighting, optics and image processing that is suitable for development in the academic environment without going through explicit 3D reconstruction and thus saving operation speed and hardware cost of the whole system.



Figure 2.15: confocal measurement used by August Ltd.

Chapter 3

Model 1: the one camera model-Homography approach

3.1 The introduction of the theoretical base of model 1

In the traditional stereo vision approach, there are unsolved shortcomings. During the inspection for a specific part of a wafer or a die, each time, we need to take at least two pictures from different view point in order to know and get the information for the 3-D information of every single point of the bump surface, which is determined by the stereo vision conception. However, sometimes we needn't to know so much information, in some inspection industrial, we just need to know whether the height of the bump meet the requirement or not, thus reject the wafer or accept it. In this model, we name it the Homgraphy approach, because we make full use the information the "Homography" matrix that tells us.

3.1.1 The objective of model 1

1) Coplanarity Check and Bad Bump Detection: Assuming MOST bumps are of the DESIRABLE HEIGHT, find the plane that contains most of the bump peaks (or a surface point near the peak), and locate bumps that are outliers to the coplanarity condition (i.e., their heights (in the direction orthogonal to the wafer plane) deviate TOO MUCH (thus need a threshold of some sort) from the desired height).

- 2) Bump Shape Examination: check the above condition for a few (not just one) sample points near the peak of each bump.
- Detect the extreme case that most bumps are coplanar but NOT of the desirable height.

3.1.2 Desires

- 1) Use of light reflections from wafer bumps themselves, not additional markers on wafer, for image registration, so as to avoid introducing additional uncertainty to the problem.
- 2) Parallel operation to examine multiple bumps at the same time (so as to be fast).
- 3) Less moving parts (so as to be fast, to reduce positioning uncertainty or the need of high-cost positioning stage, and to reduce the need of frequent positioning calibration that would require the inspection system to be paused).
- 4) Simplicity (so as to reduce regular calibration effort).

3.1.3 Some background knowledge on Homography

Homography is an image-to-image mapping induced by a planar surface in 3D: for all corresponding image positions (u_i, v_i) and (u'_i, v'_i) that are projected from the same plane,

$$[u_i, v_i, 1]^T \cong H[u'_i, v'_i, 1]^T$$
 (3.1)

where H is a 3×3 matrix called Homography that is associated with that plane, and is equality up to an arbitrary scale.

According to Fig.2.13, for example, the big "P" there is a point in 3 dimensional world-coordinate. We take two pictures at it from different view points and get the two images I and II. Little p and p' are big P's projection points respectively.

The relationship of little p and p' there can be calculated from equation

$$p \cong Hp' \tag{3.2}$$

where the Homography H is only related to the big P's 3D position in the world-coordinate. In the figure, "R" means the rotation and "t" means the translation between the two cameras. "R" can be represented into a 3×3 matrix which is related to the rotation angles $\alpha\beta\gamma$ around axis xyz; and "t" can be represented as a 3×1 matrix, whose elements are the translation vector along axis xyz. Suppose there is a way to let the peaks of all the bumps in a wafer (or die) cast bright and identifiable points in an image that is pictured as the locally brightest point in the image, by the use of parallel illumination and imaging at a particular viewing angle (Fig.3.1).

Let the plane that contains all these bump peaks in 3D be

$$px + qy + cz = 1 \tag{3.3}$$

where (x, y, z) is referenced with respect to the camera coordinate frame, and z is not equal to 0. Suppose a reference wafer whose bumps are all of desirable heights and the surface of the bumps are all uncontaminated, is available for any specimen wafer to be checked against. The reference wafer could be pictured in the above configuration, and we call the image so acquired the reference image. We call the image of the specimen wafer under the same configuration the specimen image. In the simplest case where there is only one single light source, both the reference image and the specimen image are a distribution of locally brightest image points.

In this way, the bright points on the reference wafer and those on the specimen wafer will cast two separate images under the same imaging setting as in Fig.3.1. With the bump peaks identifiable as (u_i, v_i) and (u'_i, v'_i) in the two images respectively, a number of correspondences $((u_i, v_i):(u'_i, v'_i), i=1,2,3...,n, n \ge 4)$ across the two images could be established, where n is the number of the bright peaks in the images. The Homography H that maps the positions of the distinct points from one image to the other could then be estimated from such correspondences. This Homography H would be the identity matrix if the

specimen wafer has the overall wafer position in 3D, all the bump positions, and all the bump heights exactly the same as those of the reference wafer. At the same time, the Homography H will not be the identity matrix if the geometric parameters of the specimen wafer deviate from the reference values. The more the geometric parameters of the specimen wafer deviate from the reference values, the more the Homography H would deviate from the identity matrix. In this sense, the Homography H could be a good indicator about the quality of the specimen wafer.

3.2 Model 1- "Pseudo Homography" Approach

3.2.1 The description of the configuration of model 1

The configuration of the model is shown in Fig.3.1. The parallel lighting

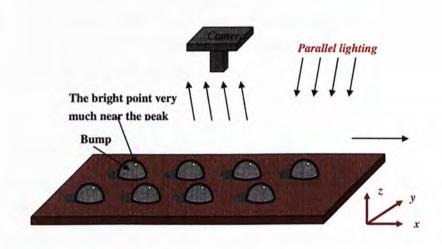


Figure 3.1: the configuration of model 1

leaves visible bright points on the area very close to the peaks of the bumps. In the camera, the image mainly concerns about these bright points. At the beginning of each inspection, feed in a golden wafer in the direction of the x axis as is shown in Fig.3.1. Use this image to be the reference image. Secondly,

feed in the wafers for inspection one by one all along the x axis direction. And each time, take one picture at the wafer with the camera. We call such images new-inspecting images. The workbench is on the x-y plane which is perpendicular to the camera above.

3.2.2 The condition of pseudo Homography

The global change of the bright points

For most cases, the precision of the feed-in system is not high enough because of the inherent mechanism system error. So, each time, on the workbench, the specimen wafers for inspection are not at the exact position as the golden (reference) wafer, they will probably have some slightly translation along the x and y axis and rotation around the z axis. And correspondingly, each time, taken from the camera, the specimen images must have some slightly differences from the reference image. Under such circumstances, though the new-inspecting images are not the same as the reference image, we still can not just reject it, for all the differences are all caused by the exterior reasons, neither by the change of the height of the bumps, nor by the layout of the bumps on the wafer. We also find that these changes happen globally on every single bump at the same time.

The individual change of the bight points

However, under other circumstances, we need to ring the alarm, when the differences between the specimen images and the reference image are caused by some original reasons from the bumps themselves, for example, the height changes or the layout changes of the bumps. This kind of change has two instances: one is that when there is bump height change, the peak positions of the bumps will have some translation along the z axis; the other is that when there is some misplacement of some individual bumps, the correspondence

bright points will change their position in the new-inspecting image individually. These two circs will both cause the individual bright point changes in the specimen images.

Additionally, from the experiences in the real-world inspection industry, we find that both the global change and the individual change always happen together, but these changes will not induce any significant differences in the specimen images. Therefore, from the images we know that, for different wafer to be inspected, images are quite alike, and much close to the reference image.

3.2.3 The formation of pseudo Homgraphy H

While, the only information we have are the series of 2D images from the specimen wafers. In the specimen images, we can only get information from those identifiable bright points. Then, how can we make use of these points to get to know the 3D information of the bumps' peaks because we have only one single image for each new-inspecting wafer? If we use the traditional stereo vision method, that will lead to impossible answer.

In stereo vision, if we want to reconstruct the 3D position from 2D views, we need to know some pairs of correspondence. Base on the condition mentioned in the last subsection, we can think about solving the problem inversely. Though the reference image and the specimen images are very near, we can assume that they just come from one object indeed, in this article, the "same wafer". With respect to model I, in the whole process of the wafer bumping inspection, the camera is fixed, only the object is changing from the reference golden wafer to every wafer to be inspected one by one. We can reverse the process to the contrary, that the object-the wafer- is settled, and camera is having some slightly rotation and translation to the same extend as the original motion of the feed in wafer for every time inspection.

In terms of visual data, a single camera observing a moving object is equivalent to a "counter-moving" camera observing a stationary object. The details about how the equivalent "counter-motion" of camera related to the object motion can be shown in the following figures. [2][3]

From the configuration of a moving object:

$$_{O2}^{C}\mathbf{T} = _{O}^{C}\mathbf{T}_{O2}^{O1}\mathbf{T} \tag{3.4}$$

From the configuration of a moving camera:

$${}_{O}^{C2}\mathbf{T} = {}_{C2}^{C1}\mathbf{T}^{-1}{}_{O}^{C}\mathbf{T} \tag{3.5}$$

Relative Transformation Equality:

$${}_{O}^{C2}\mathbf{T} = {}_{O2}^{C}\mathbf{T} \tag{3.6}$$

Thus Camera Transformation equivalent to Object Transformation is:

$${}_{C1}^{C2}\mathbf{T} = {}_{O}^{C}\mathbf{T}({}_{O2}^{O1}\mathbf{T}^{-1}){}_{O}^{C}\mathbf{T}^{-1}$$
(3.7)

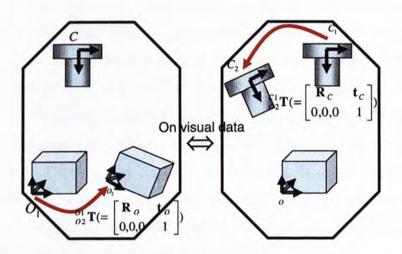


Figure 3.2: a single camera observing a moving object is equivalent to a "counter-moving" camera observing a stationary object

Thus, in model I, we can also have pairs of correspondence points. The pair of correspondence is established between the reference image and the every specimen image.

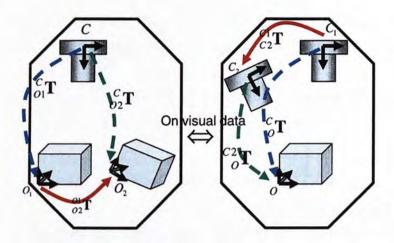


Figure 3.3: detail of the "counter-moving" camera observing a stationary object

On Camera Transformation equivalent to the Wafer Transformation

In terms of the visual data acquirable, a wafer transformation $R_{-}o$, $t_{-}o$ imaged under a fixed camera is equivalent to a counter- camera transformation $(R_{-}C, t_{-}C)$ imaging a stationary wafer. The equivalent camera transformation $(R_{-}C, t_{-}C)$ could be determined from the wafer transformation $R_{-}o$, $t_{-}o$ but the expression also involves the camera-to-reference-wafer transformation $(R_{-}Co, t_{-}Co)$. Notice that the camera-to-reference-wafer transformation and the camera-to-specimen-wafer transformation could be assumed equal, since the distance between the camera and the wafer is much larger than the magnitude of the transformation between the reference wafer and specimen wafer; we could just name the camera-to-reference-wafer transformation[4] as the camera-to-wafer transformation for simplicity.

On the values of the Wafer Transformation and the equivalent Camera

Transformation: Notice that since the wafer transformation is small, assuming that a highly repeatable positioning stage is used to position the wafers, R_o, t_o

is close to
$$\left\{ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right\}.$$
 With this fact, certain approximation could

thus be applied to various stages of the solution for the entire problem. (However, caution must be taken NOT to assume that (R_-C, t_-C) is also close to

$$\left\{ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right\}, \text{ since there is } (R_Co, t_Co) \text{ which is of great magni-}$$

tude to magnify any small deviation of R_o, t_o from $\left\{ \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \right\}.$

In Fig.3.4, bright points changes come only from the mechanism system errors, which cause the bright points on the specimen image to have some rotation and translation. Then the Homgraphy H calculated from these pairs of correspondence is only related to the so called rotation and translation[12] from the assumed two cameras. In Fig.3.5, there are only a few individual changes

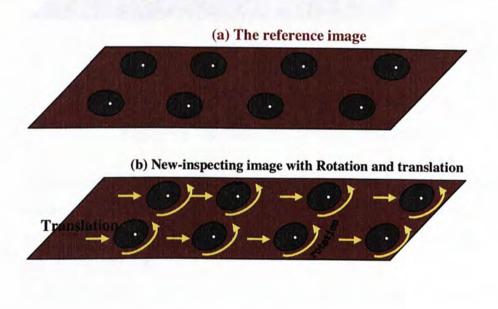


Figure 3.4: the formation of pseudo Homography

of the bright points in the specimen image. To calculate the Homography H, we first need to apply Singular Value Decomposition and Robust Estimation (please refer to the subsection 3.3: SVD and Robust Estimation) to identify

the outliers. After picking out the outliers, the left bright points have the same occupations in the image, the Homgraphy H will be an identity matrix.

In Fig.3.6, bright points changes in the new-inspecting image are found be-

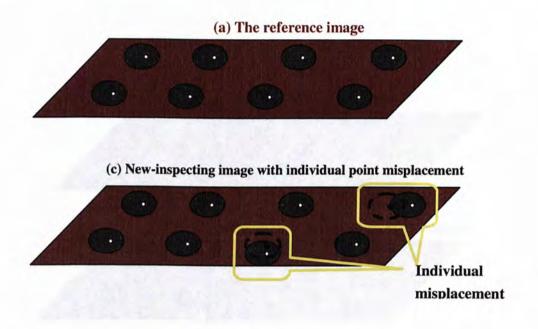


Figure 3.5: the formation of pseudo Homography

cause of some individual shift on some of the specific bumps' height. The dealing method can be the same as that for the last case (as shown in Fig.3.5), first pick out the outlier by using SVD and Robust Estimation and then estimate the Homography H.

The last possible case that can cause the global change in bright points in the new-inspecting image is shown in Fig.3.7, which is about the global height change of the wafer bumps. This case may happen when the manufacturing process (the process of making wafer bumps) has some misapplication on the bumps parameter. The Homgraphy H calculated from these pairs of correspondence will show the information about the height of the bumps to be inspected against those golden ones.

In the real inspection process, the above four instances are not separated [47]. They always happen simultaneously, but in different and unexpected proportions. Our job is just to find out how much possibility each case happened only from the Homography H we can get and draw the conclusion about whether or not these bumps are practicable for the future packaging process.

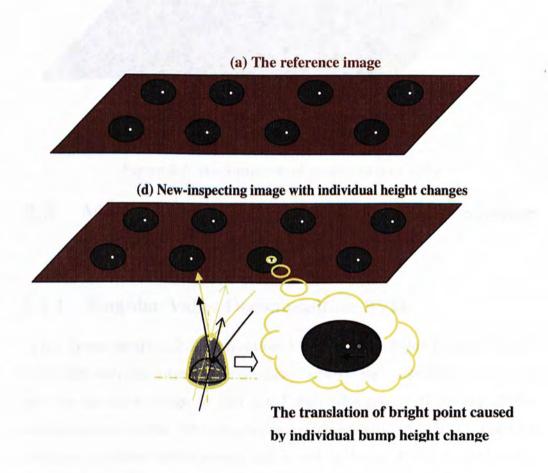


Figure 3.6: the formation of pseudo Homography

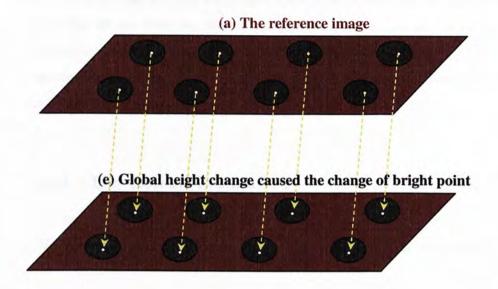


Figure 3.7: the formation of pseudo Homography

3.3 Methodology of treatment of the answer set

3.3.1 Singular Value Decomposition-SVD

In the former section.3.2.3, we mentioned the "Singular Value Decomposition" in dealing with the correspondence points between the reference golden image and the specimen image to find the Pseudo Homgraphy H among all the correspondence points. The singular value decomposition (SVD) is a powerful technique in many matrix computations and analyses. Using the SVD of a matrix in computations, rather than the original matrix has the advantage of being more robust to numerical error. Additionally the SVD exposes the geometric structure of a matrix, an important aspect of many matrix calculations. A matrix can be described as a transformation from one vector space to another. The components of the SVD quantify the resulting change between the underlying geometry of those vector spaces.

The SVD is employed in a variety of applications, from least squares problems

to solving systems of linear equations. Each of these applications exploits key properties of the SVD, its relation to the rank of a matrix and its ability to approximate matrices of a given rank. Many fundamental aspects of linear algebra rely on determining the rank of a matrix, making the SVD an important and widely used technique.

3.3.2 The Robust Estimation

The term "robust" was coined in statistics by G.E.P. Box in 1953.

Various definitions of greater or lesser mathematical rigor are possible for

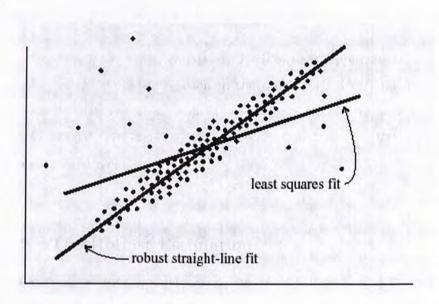


Figure 3.8: A distribution in two dimensions fitted to a straight line

the term, but in general, referring to a statistical estimator[22], it means "insensitive to small departures from the idealized assumptions for which the estimator is optimized." [22][24], The word "small" can have two different interpretations, both important: either fractionally small departures for all data points, or else fractionally large departures for a small number of data points. It is the latter interpretation, leading to the notion of outlier points,

that is generally the most stressful for statistical procedures.

Robust Estimation of the Global Mapping[65]: Homgraphy H (up to an arbitrary overall scale) could be estimated with four point correspondences. If there are many more than four point correspondences available, which is indeed the case here in wafer bump inspection, we could use robust estimation technique to estimate Homgraphy H. The procedure can is as follows:

- 1) Get the least-squared-error solution of Homgraphy H from all point correspondences;
- 2) See which particular point correspondences deviate too much, for example, exceeding a certain threshold, from the Homgraphy H mapping just estimated, and label them as outliers[25];
- 3) Use the point correspondences that are not outliers to estimate Homgraphy H again through the least-square-error method, and find outliers again as those point correspondences that do not follow this newly estimated Homgraphy H;
- 4) The above steps are iterated until the value of Homgraphy H and the set of point correspondences labeled as outliers become stable.
- 5) The robust estimation [64] gives two sets of results: the Homgraphy H that is voted positively by the majority of the point correspondences, and the set of point correspondences that are outliers to this estimated Homgraphy H. Notice also that since bumps of undesirable heights have been labeled as outliers and excluded from the estimation of H in the robust estimation process, the deviation of the robust-estimated Homgraphy H from I_3 is due only to the positioning-inaccuracy component of the camera transformation, not to the bump-height-inaccuracy component. In other words, bumps of undesirable heights do not play a role in the value of Homgraphy H that is obtained from the robust estimation [63] process.

3.3.3 Simulated experiment to test SVD[31] and Robust Estimation[24]

Here, to test the effectiveness of robust estimation and SVD to calculate the Homography H[8], we did some preliminary experiment by using some manmade Ping Pang balls.

In the following experiment, there are two ping-pang balls' pictures at first, one is the reference image, the other is the specimen image. We assume that most of the bumps here are coplanar and only one is higher than the others, which is also reasonable in the inspection process. In real world industry, on a specimen wafer, most of the bumps are coplanar, only a few of them are higher or lower or contaminated. Coplanar here means that most of the peaks of the bumps can form a plane, we call this plane the specimen peak plane. On the reference wafer, the peaks can also form such a plane, which we call the reference peak plane. The specimen peak plane can be parallel to the substrate. If the specimen peak plane is higher than the reference peak plane, it means that the bumps' height on this specimen wafer is higher than that of the golden wafer, and vise versa. If the specimen peak plane has some angle against the substrate, it means that the specimen wafer is globally tilting, which is sometimes caused by the glue. In the industrial wafer bumping inspection, there is always some glue that adhibits the specimen wafer onto the feed in system. The glue may have some thickness. While, sometimes, the glue can not be made so perfect as a cuboid, maybe like a trapezoid. That is why the specimen peak plane will have some global tilting against the substrate.

Our goal here is to find out the outlier one by using robust estimation, and calculate the Homography H for future analysis by using singular value decomposition.

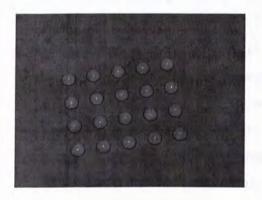
Because the brightest points have the most intensity value, I can easily find the brightest peak points by the illumination. Using all the points to be the correspondence points on the two pictures to calculate the Homography H matrix which is also very easy, but the matrix we got is not the ultimate one, because we count the outlier bump which is one-third higher than the others.

Because using the SVD of a matrix in computations rather than the original matrix has the advantage of being more robust to numerical error. So, doing the SVD at the Homography matrix we got, we can get three matrix- U, V and S[31]. The S matrix is about the singular values.

Then, I deal with the S matrix. Some of the value is so small that I regard them as errors and ignore them. So, I do the robust estimation[66] to these values. concept of the median, and that of the square error, and name a threshold, if the answer is bigger than the threshold, we know this point is much different from the others, and in this way, we can find out and remove the very small singular value. And after doing that, the dimension of the matrix is cut down. From the new V matrix, we can find the optimized Homography H. Then, we use this new Homgraphy H matrix to make product with all the points including the point of the higher bump. And we cut off the product result that is much different from the others which is caused by the higher bump.

The following pictures are the pictures before doing the SVD and robust estimation, most of the balls are coplanar and only one is higher.

The following ones are the results after doing robust estimation. The blue



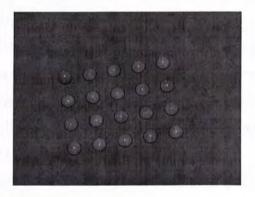
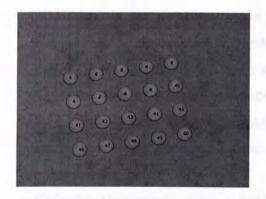


Figure 3.9: The reference image and the specimen image with an outlier cross shows the higher ball.



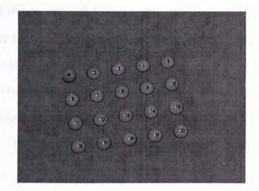


Figure 3.10: Pointing out the outlier by using robust estimation

3.3.4 the measurement of the Homography matrix answer set

After above process, we can get the optimized Homography H[9], in this step, the outlier has already been cut out. The Homography H is about the majority coplanar[18][45] peak points of the bumps on the specimen wafer.

We have mentioned before about the glue between the feed in mechanism and the specimen wafer. Though the height of the bumps on the specimen wafer will be equal to that of the bumps on the reference wafer, because of the thickness of the glue, the specimen peak plane will also probably have some global tilting against the substrate, or higher/lower than the reference peak plane.

Then, in order to avoid the misdiagnosis, (sentence the good wafer to be a bad one), we need to go on analyzing the Homgraphy H we got. We want to do the analysis from the matrix H and to see what the inherent relationship is between the changes in H and the changes of the specimen peak plane.

The detail manipulation is like this:

1) Feed in the reference wafer for the first time and get the reference image

around the peaks.

- 2) Feed in the specimen wafer for the first time and record the first specimen image about the peaks of the bumps on the specimen wafer.
- 3) Feed in the same specimen wafer for the second time, but give the wafer a slight Rotation around the z axis (refer to Fig.3.4 the configuration of model I), and some translation along x and y axis.
- 4) Iterate step 3) with the same specimen wafer for more than 9 times, and every time give the wafer some rotation and translation, which is simulating the feed-in system mechanism error. Then, record all the Homography H and make it a group.
- 5) Feed in the same specimen wafer, but this time give the wafer a slight translation along z axis and some little rotation around x and y axis (refer to Fig.3.4 the configuration of model I), which means the height of the bumps will have some change, or the glue will have some influence on the wafer's global tilting.
- 6) Iterate step 5), and record another group of Homography H.
- 7) Use the following subspace separability measurement to measure these two groups of Homography matrix. Because the two groups of Homography H are caused by different reason, we will see whether the two groups of Homography H belong to two different subspaces, and how they separate from each other.

Subspace Separability Measure

In an n-D vector space, given two vector subspaces (of equal dimension m) A and B, we need a measure that describes how separable A and B are. The following representation for subspace is used. Subspaces A and B are

represented by matrices
$$\mathbf{A}_{m \times n} = \begin{bmatrix} \mathbf{A}_1^T \\ \vdots \\ \mathbf{A}_m^T \end{bmatrix}$$
 $\mathbf{B}_{m \times n} = \begin{bmatrix} \mathbf{B}_1^T \\ \vdots \\ \mathbf{B}_m^T \end{bmatrix}$, where \mathbf{A}_i 's are

the basis vectors of Subspace A, and \mathbf{B}_i 's are basis vectors[54] of Subspace B. Suppose $A \cap B$ is the subspace at the intersection[53] of A and B. If A and B are of dimension m, $A \cap B$ is of dimension (m-1).

Then the subspace [54] a that is contained in A but orthogonal to $A \cap B$ is

one-dimensional. The subspace b that is contained in B but orthogonal to $A \cap B$ is also one-dimensional.

Suppose a and b are vectors of arbitrary nonzero magnitude in Subspaces a and b. A measure on how separable A and B could be:

$$d(A, B) = \frac{\|\underline{\mathbf{a}} \times \underline{\mathbf{b}}\|}{\|\mathbf{a}\| \|\mathbf{b}\|} \in [0, 1] \quad \begin{cases} 0 & \text{exactly overlapping} \\ 1 & \text{orthogonal} \end{cases}$$
(3.8)

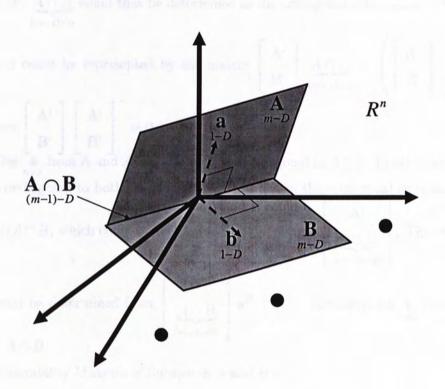


Figure 3.11: the separation of subspace

Steps of Arriving at the Separability[53] Measure are shown below:

- 1. Get $\underbrace{A \cap B}_{(m-1)\times n}$ a) Get the matrix $\underbrace{A'}_{(n-m)\times n}$ that represents the orthogonal subspace A' of A: $\underbrace{A'}_{(n-m)\times n} = (\mathbf{A}^{\perp})^{\mathrm{T}}$, where \mathbf{A}^{\perp} represents the Null space of A and could be
- determined from $\mathbf{A}_{m \times n} \mathbf{A}_{n \times (n-m)}^{\perp} = \mathbf{0}$ b) Similarly, get the matrix $\mathbf{B}'_{(n-m) \times n}$ that represents the orthogonal subspace B' of B.
- c) The subspace $A \cap B$ at the intersection of A and B is orthogonal to both A' and B'. $\underbrace{A \cap B}_{(m-1) \times n}$ could thus be determined as the orthogonal subspace of $A' \cup B'$

which could be represented by the matrix $\begin{bmatrix} \mathbf{A}' \\ \mathbf{B}' \end{bmatrix} \underbrace{\underbrace{A \cap B}_{(m-1) \times n}} = \left(\begin{bmatrix} \mathbf{A}' \\ \mathbf{B}' \end{bmatrix} \right)$,

where $\begin{bmatrix} A' \\ B' \end{bmatrix} \begin{bmatrix} A' \\ B' \end{bmatrix}^{\perp} = 0$

2. Get $\underset{n \neq 1}{\mathbf{a}}$ from A and $A \cap B$ a is in A but orthogonal to $A \cap B$. In other words, a is orthogonal to both A' and $A \cap B$. Thus a is in the orthogonal subspace of

a is orthogonal to both A and A' $A' \cup (A \cap B)$ which could be represented by the matrix $A \cap B \cup A'$. Therefore,

a could be determined from $\begin{bmatrix} A' \\ \underbrace{A \cap B}_{n \times 1} \end{bmatrix} \mathbf{a}^{\mathrm{T}} = \mathbf{0}.$ 3. Similarly, get $\mathbf{b}_{n \times 1}$ from B and $A \cap B$.

4. Separability Measure of Subspaces A and B is:

$$d(A, B) = \frac{\|\underline{\mathbf{a}} \times \underline{\mathbf{b}}\|}{\|\mathbf{a}\| \|\mathbf{b}\|} \in [0, 1] \quad \begin{cases} 0 & \text{exactly overlapping} \\ 1 & \text{orthogonal} \end{cases}$$
(3.9)

The union[55] of two subspace

Assume that there are two subspaces A and B, A and B can be represented as $\begin{bmatrix} \alpha_1 & \alpha_2 \end{bmatrix}$ and $\begin{bmatrix} \beta_1 & \beta_2 \end{bmatrix}$ respectively, where α , β are all n by one vectors. To get to know the union of the two subspaces is just to know the span of $\alpha_1,\alpha_2,\beta_1,\ \beta_2.$

We know that the union of the two subspaces is a linearization[56] representation of the subspaces A and B, so span[54] $(\alpha_1,\alpha_2,\beta_1, \beta_2)$ can be written as below:

$$span = \iota_1 \alpha_1 + \iota_2 \alpha_2 + \iota_3 \beta_1 + \iota_4 \beta_2 \tag{3.10}$$

Mathematically, to get the coefficients in this equation, there is an easy method, that is doing the SVD of Ψ , where:

$$\Psi = \begin{bmatrix} \alpha_1 & \alpha_2 & \beta_1 & \beta_2 \end{bmatrix} \tag{3.11}$$

Then, after simplification, we know the union of the two subspaces, which is:

$$\Gamma = \begin{bmatrix} \gamma_1 & \gamma_2 & \gamma_3 & \gamma_4 \end{bmatrix} \tag{3.12}$$

Now, there is a new subspace here represented as χ , which is also a n by one vector. We want to see that how much it belongs to the two subspaces A and B, which means that we need to measure the angles between χ and A, B respectively. But before we measure the angles, we first should project the vector onto the union to make sure that the angle between χ and A, B, is the sum of the angles between χ and A and χ and B. For any arbitrary vector χ in the 9-dimensional space[57], it can be represented as:

$$\chi = \lambda_1 \gamma_1 + \lambda_2 \gamma_2 + \lambda_3 \gamma_3 + \lambda_4 \gamma_4 + \lambda_\perp \gamma_\perp \tag{3.13}$$

Where γ_{\perp} represents the orthogonal subspace[57] of the union Γ , Then, we apply dot product to ignore the part of γ_{\perp} , for the projection of vector χ on the union Γ is just the part of:

$$\chi_p = \lambda_1 \gamma_1 + \lambda_2 \gamma_2 + \lambda_3 \gamma_3 + \lambda_4 \gamma_4 \tag{3.14}$$

named as χ_p :

$$\chi \cdot \gamma_1 = \lambda_1 \gamma_1 \cdot \gamma_1 + \lambda_2 \gamma_2 \cdot \gamma_1 + \lambda_3 \gamma_3 \cdot \gamma_1 + \lambda_4 \gamma_4 \cdot \gamma_1 + \lambda_\perp \gamma_\perp \cdot \gamma_1
\Rightarrow \chi \cdot \gamma_1 = \lambda_1 \gamma_1 \cdot \gamma_1 + \lambda_2 \gamma_2 \cdot \gamma_1 + \lambda_3 \gamma_3 \cdot \gamma_1 + \lambda_4 \gamma_4 \cdot \gamma_1$$
(3.15)

And the same:

$$\chi \cdot \gamma_2 = \lambda_1 \gamma_1 \cdot \gamma_2 + \lambda_2 \gamma_2 \cdot \gamma_2 + \lambda_3 \gamma_3 \cdot \gamma_2 + \lambda_4 \gamma_4 \cdot \gamma_2 \tag{3.16}$$

$$\chi \cdot \gamma_3 = \lambda_1 \gamma_1 \cdot \gamma_3 + \lambda_2 \gamma_2 \cdot \gamma_3 + \lambda_3 \gamma_3 \cdot \gamma_3 + \lambda_4 \gamma_4 \cdot \gamma_3 \tag{3.17}$$

$$\chi \cdot \gamma_4 = \lambda_1 \gamma_1 \cdot \gamma_4 + \lambda_2 \gamma_2 \cdot \gamma_4 + \lambda_3 \gamma_3 \cdot \gamma_4 + \lambda_4 \gamma_4 \cdot \gamma_4 \tag{3.18}$$

Assume the matrix is:

$$\mathbf{M} = \begin{bmatrix} \gamma_1 \cdot \gamma_1 & \gamma_2 \cdot \gamma_1 & \gamma_3 \cdot \gamma_1 & \gamma_4 \cdot \gamma_1 \\ \gamma_1 \cdot \gamma_2 & \gamma_2 \cdot \gamma_2 & \gamma_3 \cdot \gamma_2 & \gamma_4 \cdot \gamma_2 \\ \gamma_1 \cdot \gamma_3 & \gamma_2 \cdot \gamma_3 & \gamma_3 \cdot \gamma_3 & \gamma_4 \cdot \gamma_3 \\ \gamma_1 \cdot \gamma_4 & \gamma_2 \cdot \gamma_4 & \gamma_3 \cdot \gamma_4 & \gamma_4 \cdot \gamma_4 \end{bmatrix}$$
(3.19)

Then,

$$M \cdot \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{pmatrix} = \chi \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_4 \end{pmatrix}$$

$$\Rightarrow \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \end{pmatrix} = M^{-1} \cdot \chi \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \\ \gamma_4 \end{pmatrix}$$

$$(3.20)$$

In this way we can get the coefficients of the

$$\chi_p = \lambda_1 \gamma_1 + \lambda_2 \gamma_2 + \lambda_3 \gamma_3 + \lambda_4 \gamma_4 \tag{3.21}$$

and know the χ_p .

3.4 Preliminary experiment about model 1

Preliminary experiments show that isolated bumps with peaks not coplanar with the peaks of other bumps in the same wafer could be easily identified by using robust estimation technique. It seems that their presence does not complicate the problem by too much.

By this time, we want to separate two cases:

- One is about the changes of the bright points on the specimen image are caused by the feed-in mechanism system error, which we should ignore;
- 2) The other is about the changes of the bright points on the specimen image are caused by the bump height change, which we will find out.

Sometimes, these two cases will happen together, so we need to find out what the information that the ultimate Homography H tells us, that is to say, from the Homography, we need to judge how much the changes rely on the system error, and how much that does on the bump's height change.

The procedure can be described by the followed flow chart:

First give the reference wafer a few rotation around the z axis, and slight translation along x and y axis [49](according to figure) to simulate the feed-in system error and record one group of Homography matrix; then, give the reference wafer a few rotation around x and y axis and slight translation along z axis and record the other group of Homography to simulate the bump Height change. These two groups will be the reference for the future decomposition step.

Then, we apply the specimen wafer, following the same procedure as mentioned in page 35. We use some synthetic data in the preliminary experiment to test this idea. And the intrinsic parameters[6] for the cameras are:

There are 30 correspondence points on the specimen wafer. The relationship between the points in the world coordinate and the world in the image coordinate[7] is as shown below:

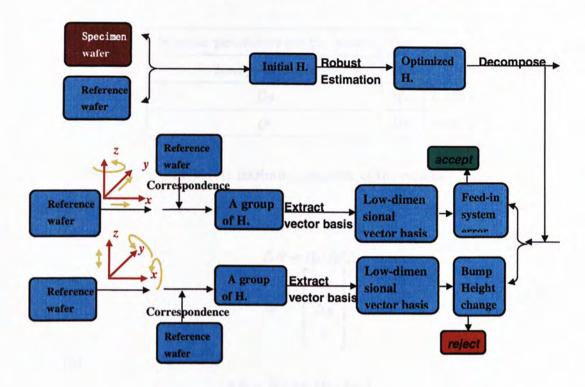


Figure 3.12: Flow chart of inspection procedure by model 1

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{f}{Du} & 0 & 0 \\ 0 & \frac{f}{Dv} & 0 \\ 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} * \begin{bmatrix} R & t \\ 0 & 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} Xw \\ Yw \\ Zw \\ 1 \end{bmatrix}$$
(3.22)

Assume the rotation and translation matrix is as below: When there is only rotation around z axis and only translation along x axis and y axis which we called the condition a, when there is only rotation around x axis and y axis and translation along z axis which we called the condition b, the change in rotation matrix and translation matrix is represented as below.

Intrinsic parameters for the camers	Unit	Value
Lens focal length	Mm	50
Du	Mm	0.006
Dv	Mm	0.006

Table 3.1: Intrinsic parameter of the camera

(a)
$$\Delta R = R_0(\Delta \theta_z)$$

$$\Delta t = \begin{bmatrix} \Delta x \\ \Delta y \\ 0 \end{bmatrix}$$
(3.23)

(b)
$$\Delta R = R_0(\Delta \theta_x) R_0(\Delta \theta_y)$$

$$\Delta t = \begin{bmatrix} 0 \\ 0 \\ \Delta z \end{bmatrix}$$

$$\Delta Rt = \begin{bmatrix} \Delta R & \Delta t \\ 0 & 0 & 0 & 1 \end{bmatrix}$$
(3.24)

The points after rotation and translation changed into:

$$\begin{bmatrix} u' \\ v' \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{f}{Du} & 0 & 0 \\ 0 & \frac{f}{Dv} & 0 \\ 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} * \begin{bmatrix} R & t \\ 0 & 0 & 0 & 1 \end{bmatrix} * \Delta Rt * \begin{bmatrix} Xw \\ Yw \\ Zw \\ 1 \end{bmatrix}$$
(3.25)

Then, calculate the Homography between the two points:

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \text{ and } \begin{bmatrix} u' \\ v' \\ 1 \end{bmatrix} \text{ Then, we get two groups of Homography matrix.}$$

After doing the Singular value decomposition of the Homography matrix under

two different conditions, we can get the vector basis of the two subspaces. There are 30 synthetic bumps on the wafer, we focus on the brightest points of each bump, so there are 30 points on each wafer which are the correspondence points we concerned. The Homography matrix is a 9 by 30 matrix. The vector basis is 9 by 2. So the dimension of the two subspaces is both 2, and the angle[58] between the two subspaces is nearly 90 degree. The answer is gotten from the synthetic data, some of which is shown in the figure above, and verified by the experiment on some pattern pictures.

The following are some raw data in the experiments.

(a) synthetic data (simulate the bright points)

$$x1 = [-80, 0, 80, -40, 40, -60, 60, -20, 20, -80, 80, -20, 20, -60, 60, ...];$$

 $y1 = [80, 80, 80, 60, 60, 40, 40, 20, 20, 0, 0, -20, -20, -40, -40, ...]$

The two vector basis called basis_u1 and basis_u2

$$basis_u1 = \begin{bmatrix} -0.0000, 0.0000 \\ 0.0000, 0.0000 \\ -0.4593, 0.4667 \\ 0.0268, -0.0512 \\ 0.5252, -0.5406 \\ 0.7158, 0.6981 \\ 0.0104, -0.0028 \\ -0.0000, -0.0000 \\ 0.0000, 0.0000 \end{bmatrix}$$

and

$$basis_u2 = \begin{bmatrix} -0.0000, 0.0000 \\ 0.0000, -0.0000 \\ -0.0939, 0.0911 \\ -0.7049, 0.6971 \\ 0.4801, 0.4911 \\ -0.0517, -0.0500 \\ 0.5110, 0.5119 \\ 0.0000, 0.0000 \\ -0.0000, -0.0000 \end{bmatrix}$$
 in the two subspaces is 88.9279 degree.

The angle between the two subspaces is 88.9279 degree. These raw data can somewhat explain that the two kinds of subspace can be separable, and the angle between them is near 90 degree.

3.5 Problems unsolved

However, there are still some problems this model cannot solve. We mentioned about the glue[49] between the specimen wafer and the feed-in system, the shape of the glue is not standard, sometimes it is the glue but not the height of the bump itself that causes the global tilting or global translation along the z axis about the bright point plane. Once this happens, the inspection system will draw a wrong conclusion, thus cause the mis-judgment.

So, to think about solving this problem and improve the system, we get some modification about model 1 and advert model 2- the two camera system.

Chapter 4

Model 2: the two camera model-Biplanar Disparity approach

4.1 Theoretical Background

In the last step we can solve the following problems:

- 1) The global transformation caused the translation along x and y axis, and rotation around the z axis which are all planar changes about the bright points in the image which is the change within the wafer plane;
- 2) The global parameter changes about the wafer plane, which is the changes against the wafer plane, with the peak on the images changes together;
- 3) For the former two cases, the changes in the Homography matrix are not linear, but locally, we can assume them to be linear after some linearization mathematically.
- 4) For each specimen wafer, the calculated Homography can be projected onto the union space of subspaces induced from the 1 and 2 cases above, and then decomposed with the union space onto the two subspaces. After simplification, we know that the last two cases 1 and 2 can be obviously separable.

4.1.1 the linearization of Homography matrix changes

There are two kinds of changes in the thesis about the bright points on the specimen wafer. One is about the specimen peak plane is higher or lower or tilting globally. The other is about the two dimensional changes on the specimen peak plane, the translation along the x and y axis and the rotation around the z axis which is also globally. We can see from the following equation that after linearization, the Homography matrix can be written into the old one plus three constant matrix times the variable. With respect to the former knowledge yuanding's paperwe can write the Homography in the following way:

$$H \cong \begin{bmatrix} R_{11} + t_x p & R_{12} + t_x q & R_{13} + t_x c \\ R_{21} + t_y p & R_{22} + t_y q & R_{23} + t_y c \\ R_{31} + t_z p & R_{32} + t_z q & R_{33} + t_z c \end{bmatrix}$$
(4.1)

where px + qy + cz = 1 and $c \neq 0$ is the wafer plane equation.

1) For the first case, as simulating the feed-in mechanism system error[50]:

$$R = R_x * R_y * R_z$$

$$R_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & \sin \alpha \\ 0 & -\sin \alpha & \cos \alpha \end{bmatrix}$$

$$R_z = \begin{bmatrix} \cos \gamma & \sin \gamma & 0 \\ -\sin \gamma & \cos \gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Assume that:

$$\sin \alpha = a_1$$
 $\cos \alpha = a_2$ $so,$ $a_1^2 + a_2^2 = 1$
 $\sin \beta = b_1$ $\cos \beta = b_2$ $so,$ $b_1^2 + b_2^2 = 1$ (4.2)
 $\sin \gamma = k_1$ $\cos \gamma = k_2$ $so,$ $k_1^2 + k_2^2 = 1$

Also:

$$t'_{x} = t_{x} + \Delta x$$

$$t'_{y} = t_{y} + \Delta y$$

$$R'_{z} = R_{z} + \Delta R_{z}$$

$$R'_{z} = f(k'_{1}, k'_{2})$$

$$(4.3)$$

$$k'_{1} = k_{1} + \Delta k_{1}$$
 $(\Delta k_{1} = \Delta k)$ (4.4)

$$k'_{2} = k_{2} + \Delta k_{2}$$

$$\Delta k_{2} = \frac{k_{1} + k_{2}}{k_{1} - k_{2}} \Delta k_{1}$$

All the others remain constant, so we can get:

$$H' \cong H + \begin{bmatrix} p & q & c \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \Delta x + \begin{bmatrix} 0 & 0 & 0 \\ p & q & c \\ 0 & 0 & 0 \end{bmatrix} \Delta y$$

$$+ \begin{bmatrix} \frac{k_1 + k_2}{k_1 - k_2} b_2 & b_2 & 0 \\ -a_2 + \frac{k_1 + k_2}{k_1 - k_2} a_1 b_1 & \frac{k_1 + k_2}{k_1 - k_2} a_2 + a_1 b_1 & 0 \\ a_1 + \frac{k_1 + k_2}{k_1 - k_2} a_2 b_1 & -\frac{k_1 + k_2}{k_1 - k_2} a_1 + a_2 b_1 & 0 \end{bmatrix} \Delta k$$

$$(4.5)$$

2) For the second case, with the same method, we can get the Homography after the specimen peak plane changes:

$$t'_{z} = t_{z} + \Delta z$$

$$R'_{x} = R_{x} + \Delta R_{x}$$

$$R'_{y} = R_{y} + \Delta R_{y}$$

$$R'_{y} = f(b'_{1}, b'_{2})$$

$$(4.6)$$

$$a'_{1} = a_{1} + \Delta a_{1}$$
 $(\Delta a_{1} = \Delta a)$ $a'_{2} = a_{2} + \Delta a_{2}$ (4.7)
 $b'_{1} = b_{1} + \Delta b_{1}$ $(\Delta b_{1} = \Delta b)$ $b'_{2} = b_{2} + \Delta b_{2}$

$$H' \cong H + \begin{bmatrix} \frac{b_1 + b_2}{b_1 - b_2} k_2 & \frac{b_1 + b_2}{b_1 - b_2} k_1 & -1 \\ a_1 k_2 & a_1 k_1 & \frac{b_1 + b_2}{b_1 - b_2} a_1 \\ a_2 k_2 & a_2 k_1 & \frac{b_1 + b_2}{b_1 - b_2} a_2 \end{bmatrix} \Delta b + \begin{bmatrix} 0 & 0 & 0 \\ b_1 k_2 - \frac{a_1 + a_2}{a_1 - a_2} k_1 & b_1 k_1 + \frac{a_1 + a_2}{a_1 - a_2} k_2 & b_2 \\ k_1 + \frac{a_1 + a_2}{a_1 - a_2} b_1 k_2 & -k_2 + \frac{a_1 + a_2}{a_1 - a_2} b_1 k_1 & \frac{a_1 + a_2}{a_1 - a_2} b_2 \end{bmatrix} \Delta a + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ p & q & c \end{bmatrix} \Delta z$$

$$(4.8)$$

So, after the linearization[58], we can see that no matter what kind of changes, in some local area, the new Homography is just an approximated linear translation of the old Homography with three variables.

4.1.2 Problem Nature

To summarize, all the above listed four questions can be solved by now, but for real inspection of wafer bump, the global changes caused by the glue really exist, as well as the individual change of the bumps. To identify the little translation along x and y axis against the little change in the z direction which is the height change of the bump itself, we need to know not only the plane about the peaks, but also the points about the bump bottoms.

We now understand the followings: Bump surface has a mirror-like reflective property, and:

- 1) In 3D inspection, the bumps should be allowed to be displaced twodimensionally on the wafer plane relative to one another. In other words, 3D inspection should not be about the 3D positions of individual bumps, but about the overall planes Π_t and Π_b that contain the bump peaks and the bump bottoms respectively.
- 2) 3D inspection of bumps should not be about the distance of the bump peaks from a camera, but about the heights of the bumps, i.e., about the distance of bumps' peaks from their bottoms. Together with the above, the 3D inspection problem should be about the difference of the 2 planes Π_t and Π_b
- 3) Isolated bumps whose heights deviate too much from the average height, i.e., those whose peaks are too distant from the plane Π_t , should be identified.

Here we introduce a new approach which is also based on the Homography background knowledge[11]. The imaging is setup in the Fig.1.1.

4.1.3 Imaging system setup

In the configuration of the lighting, we add an illumination in model II. The red camera called camera 1 takes images perpendicular to the substrate of the specimen wafer, and the red illumination is in the same direction; the blue lighting and the side camera (camera 2) are symmetrically placed forming an angle, with the optical axis of camera1 the angle bisector. In this way, the bright points in image1 and bright points in image 2 are 3D the same point-the peak point of the bumps on the specimen wafer.

4.1.4 System Calibration[13]

To make sure the system is authentic, before the inspection, we should do the calibration once for all, the procedure is as follows:

- 1) Put a mirror which is of the same thickness as the reference wafer, to the position of the wafer.
- 2) Position Camera 1 so that its optical axis passes through the center of the mirror, and is as orthogonal to the mirror as possible.
- 3) Position the Red parallel light source so that its parallel light covers the mirror, and is also as orthogonal to the mirror as possible.
- 4) Move the Red light source left and right, up and down until the red reflections perceived by Camera 1 is the "brightest". That way we are sure that the angle bisector of the Red light source and the optical axis of Camera 1 is approximately parallel to the surface normal of the mirror.
- 5) Position Camera 2 so that its optical axis passes through the center of the mirror, and at an angle θ from the optical axis of Camera 1 with respect to the mirror. Suppose the plane that contains the optical axes of both cameras is Π .
- 6) Position the Blue light source so that its parallel light covers the mirror, is contained in Plane Π , and is at an angle approximately θ from the optical axis

- of Camera 1 (with respect to the mirror) on the opposite side of Camera 2.
- 7) Move the Blue light source left and right, up and down until the blue reflections perceived by Camera 2 is the "brightest". That way we are sure that the angle bisector of the Blue light source and the optical axis of Camera 2 is also approximately parallel to the surface normal of the mirror.

An imaging setup that allows the planes Π_t and Π_b be accessible in two images I and I' respectively, is the following:

- 1) Exploiting the mirror-like reflectivity of the bumps, 2 cameras and 2 parallel light sources of different colors, if properly positioned in 3D (so that in 3D the red dot and the blue dot coincide) as illustrated in Figure 1, and coupled with the use of the epipolar constraint[10], which is captured by a 33 matrix F named the Fundamental matrix, could allow image point correspondences over the bump peaks, and over the bump bottoms, to be available over a number of bumps in the wafer.
- 2) Such 2 sets of correspondences would allow the two Homographies $\mathbf{H_t}$ and $\mathbf{H_b}$, which are for the planes Π_t and Π_b respectively, to be estimated through the use of the basic function $\begin{bmatrix} \mathbf{p'} \\ 1 \end{bmatrix} \cong \mathbf{H_t} \begin{bmatrix} \mathbf{p} \\ 1 \end{bmatrix}$.
- 3) Isolated bumps whose heights deviate too much from the majority value could be identified using SVD in the estimation of \mathbf{H}_{t} and \mathbf{H}_{b} . In other words, Problem (3) above is solved.
- 4) The use of Homography makes the whole system be sensitive not to the individual positions of the bumps but the overall positions of planes Π_t and Π_b only. In other words, Problem (1) above is solved.
- 5) The remaining problem to solve is Problem(2). On it the key question is: How could we come up with an invariant about the difference of planes Π_t and $\Pi_b[14]$, which should be more or less preserved (to the limit of a threshold) across wafers?

4.2 Methodology

4.2.1 Invariance measure

Homography Estimation

The outliers whose heights are too high or too low should be detected before estimating the Homographies by the method of Robust Estimation[26] as mentioned in last subsection.

Then, as shown in figure 4.1, two sets of correspondences would allow the two Homographies \mathbf{H}_t and \mathbf{H}_b , which are for the planes Π_t (the plane which contains most of the peak points of the bumps) and Π_b respectively, to be estimated through the use of the property:

$$\begin{bmatrix} \mathbf{p}_i^{\prime t} \\ 1 \end{bmatrix} \cong \mathbf{H}_t \begin{bmatrix} \mathbf{p}_i^t \\ 1 \end{bmatrix} \tag{4.9}$$

for all the i;

$$\begin{bmatrix} \mathbf{p}_{j}^{\prime b} \\ 1 \end{bmatrix} \cong \mathbf{H}_{b} \begin{bmatrix} \mathbf{p}_{j}^{b} \\ 1 \end{bmatrix} \tag{4.10}$$

for all the j;

Proof of Homography expression

Suppose we have access to 2 images I and I' of the same plane Π , and the camera intrinsic parameter matrices[7] A and A' of the 2 cameras. The Homography H induced by Π to the 2 images is related to the plane parameters $(\hat{\mathbf{n}}, d)$ by:

$$\mathbf{H} \cong \mathbf{A}' \mathbf{R}^{-1} (\mathbf{I} - \mathbf{t} \frac{\hat{\mathbf{n}}^{\mathrm{T}}}{d}) \mathbf{A}^{-1}$$
 (4.11)

Where $\hat{\mathbf{n}}$ is the unit surface normal vector of plane Π with reference to the camera coordinate frame of I, and d is the perpendicular distance of the plane to the camera center of I.

Proof:
$$\mathbf{H} \cong \mathbf{A}'\mathbf{R}^{-1}(\mathbf{I} - \mathbf{t}\frac{\hat{\mathbf{n}}^{\mathrm{T}}}{d})\mathbf{A}^{-1}$$

Let \mathbf{P} be the position vector of any 3D point on Π , with respect to the camera coordinate frame of I. Let \mathbf{P}' be the corresponding position vector from the camera coordinate frame of I'.

Then $\hat{\mathbf{n}} \cdot (\mathbf{P} - \mathbf{P}_0) = 0$, where \mathbf{P}_0 is the position of a fixed point on Π . The above could be simplified to

$$\hat{\mathbf{n}} \cdot \mathbf{P} = \hat{\mathbf{n}} \cdot \mathbf{P}_0 = d \tag{4.12}$$

We also have:

$$\mathbf{P} = \mathbf{RP'} + \mathbf{t} \tag{4.13}$$

Equation 4.12 and Equation 4.13 together imply

$$\mathbf{P}' = \mathbf{R}^{-1}\mathbf{P} - \mathbf{R}^{-1}\mathbf{t}(\frac{1}{d}\mathbf{n}^{\mathrm{T}}\mathbf{P})$$
 (4.14)

or simply:

$$\mathbf{P}' = \mathbf{R}^{-1}(\mathbf{I} - \mathbf{t}\frac{\mathbf{n}^{\mathrm{T}}}{d})\mathbf{P}$$
 (4.15)

Since the corresponding [27] image positions $(\mathbf{p}, \mathbf{p}')$ in I and I' have the properties:

$$\mathbf{p} = (x, y, 1)^{\mathrm{T}} \cong \mathbf{A}[\mathbf{I}, 0][\mathbf{P}^{\mathrm{T}}, 1]^{\mathrm{T}} \cong \mathbf{A}\mathbf{I}\mathbf{P} \cong \mathbf{A}\mathbf{P}$$
(4.16)

and

$$\mathbf{p}' = (x', y', 1)^{\mathrm{T}} \cong \mathbf{A}' \mathbf{P}' \tag{4.17}$$

the above means

$$\mathbf{p}' \cong [\mathbf{A}'\mathbf{R}^{-1}(\mathbf{I} - \mathbf{t}\frac{\mathbf{n}}{d}^{\mathrm{T}})\mathbf{A}^{-1}]\mathbf{p}$$
 (4.18)

The expression of H is thus proved.

Equation 4.11 for a normalized $\hat{\mathbf{H}}$ could be written as:

$$\hat{\mathbf{H}} = w\mathbf{A}'\mathbf{R}^{-1}(\mathbf{I} - \mathbf{t}\frac{\hat{\mathbf{n}}}{d}^{\mathrm{T}})\mathbf{A}^{-1}$$
(4.19)

for some unknown w. The equation could be simplified to:

$$[\mathbf{t}]_{\times} \mathbf{R} \mathbf{A}^{\prime - 1} \hat{\mathbf{H}} = w [\mathbf{t}]_{\times} \mathbf{A}^{-1}$$
(4.20)

which could be expressed as:

$$\mathbf{F}\hat{\mathbf{H}} = w(\mathbf{A}^{-1})^{\mathrm{T}} [\mathbf{t}] \wedge \mathbf{A}^{-1}$$
(4.21)

where,

$$\mathbf{F} = (\mathbf{A}^{-1})^{\mathrm{T}} \left[\mathbf{t} \right]_{\times} \mathbf{R} \mathbf{A}'^{-1} \tag{4.22}$$

is the fundamental matrix of the camera pair. In other words, given unit-norm $\hat{\mathbf{H}}$ and unit-norm[59] $\hat{\mathbf{F}}$, the unknown scale w could be determined in terms of a particular scalar from:

$$\alpha \hat{\mathbf{F}} \hat{\mathbf{H}} = w(\mathbf{A}^{-1})^{\mathrm{T}} [\mathbf{t}]_{\star} \mathbf{A}^{-1}$$
(4.23)

Where α is a constant related to the values of A,A',R, and t., not Π_t and Π_b Suppose $\hat{\mathbf{H}}_t$ and $\hat{\mathbf{H}}_b$ are the unit-norm \mathbf{H}_t and \mathbf{H}_b for planes Π_t and Π_b respectively. Suppose w_t and w_b are the unknown scales in Equation 4.9 and Equation 4.10 for $\hat{\mathbf{H}}_t$ and $\hat{\mathbf{H}}_b$, $\frac{w_t}{\alpha}$ and $\frac{w_b}{\alpha}$, for some α , could be determined respectively from $\hat{\mathbf{H}}_t$, $\hat{\mathbf{H}}_b$, and $\hat{\mathbf{F}}$ through the use of Equation 4.23. given $\hat{\mathbf{F}}$, $\hat{\mathbf{H}}$, \mathbf{A} , \mathbf{t} . Because of noise etc., the left and right sides of Equation 4.23 might not be equal up to a scale as expected, and that makes the estimation of $\frac{w}{\alpha}$ tricky.

Least-Squares-Error Solution[62] for $\frac{w}{a}$

First normalize $\hat{\mathbf{F}}$ and $\hat{\mathbf{H}}$ so that they are of unit norm. Then determine $\hat{\mathbf{F}}\hat{\mathbf{H}}$ and $(\mathbf{A}^{-1})^{\mathrm{T}}[\mathbf{t}]_{\times}\mathbf{A}^{-1}$. Denote $\frac{w}{\alpha}$ as β , $\hat{\mathbf{F}}\hat{\mathbf{H}}$ as matrix \mathbf{G} , and $(\mathbf{A}^{-1})^{\mathrm{T}}[\mathbf{t}]_{\times}\mathbf{A}^{-1}$ as matrix \mathbf{B} .

Under the least-squares-error estimation criterion, estimating β is about finding the $\bar{\beta}$ such that:

$$E(\beta)|_{\beta=\bar{\beta}} = \|\mathbf{G} - \beta\mathbf{B}\|^2|_{\beta=\bar{\beta}}$$
(4.24)

is minimum.

 $E(\beta)$ could be written as:

$$E(\beta) = \sum_{ij} (G_{ij} - \beta B_{ij})^2 = \sum_{ij} (G_{ij}^2 - 2\beta G_{ij} B_{ij} + \beta^2 B_{ij}^2)$$

= $\sum_{ij} G_{ij}^2 - 2\beta \sum_{ij} G_{ij} B_{ij} + \beta^2 \sum_{ij} B_{ij}^2$ (4.25)

Thus minimizing $E(\beta)$ is equivalent to minimizing:

$$E'(\beta) = -2\beta(\mathbf{G} \cdot \mathbf{B}) + \beta^2 \|\mathbf{B}\|^2$$
(4.26)

where
$$\mathbf{G} \cdot \mathbf{B} = \sum_{ij} G_{ij} B_{ij}$$
, $\|\mathbf{B}\|^2 = \sum_{ij} B_{ij}^2$.
At $\beta = \bar{\beta}$, $\frac{dE'(\beta)}{d\beta}\Big|_{\beta = \bar{\beta}} = 0$,
which says:

$$-2(\mathbf{G} \cdot \mathbf{B}) + 2\bar{\beta} \|\mathbf{B}\|^2 = 0 \tag{4.27}$$

Thus $\bar{\beta}$ could be estimated as:

$$\bar{\beta} = \frac{\mathbf{G} \cdot \mathbf{B}}{\|\mathbf{B}\|^2} \tag{4.28}$$

The error of the above estimation could be determined as $E(\bar{\beta})$, normalized by the norms of G and B.

Estimation error measure:

$$\frac{E(\beta)|_{\beta = \bar{\beta}}}{\|\mathbf{G}\| \|\mathbf{B}\|} = \frac{\|\mathbf{G} - \beta\mathbf{B}\|^2|_{\beta = \bar{\beta}}}{\|\mathbf{G}\| \|\mathbf{B}\|}$$
(4.29)

A measure that carries more geometric meaning is however the normalized cross product of G and B: Sine of angle between G and B is:

$$\frac{\|\mathbf{G} \times \mathbf{B}\|}{\|\mathbf{G}\| \|\mathbf{B}\|} \in [0, 1] \tag{4.30}$$

Ideally, both measures should be zero. In practice, they should be a positive value close to zero. If they are not close to zero, the estimation process has a problem.

Biplanar Disparity Matrix D

Since,

$$\mathbf{\hat{H}_t} = w_{\mathrm{t}} \mathbf{A}' \mathbf{R}^{-1} (\mathbf{I} - \mathbf{t} \frac{\mathbf{\hat{n}_t}^{\mathrm{T}}}{d_{\mathrm{t}}}) \mathbf{A}^{-1}$$

and

$$\mathbf{\hat{H}_b} = w_\mathrm{b} \mathbf{A}' \mathbf{R}^{-1} (\mathbf{I} - \mathbf{t} \frac{\mathbf{\hat{n}_b^T}}{d_\mathrm{b}}) \mathbf{A}^{-1}$$

we have:

$$(\hat{\mathbf{H}}_{t}/\frac{w_{t}}{\alpha} - \hat{\mathbf{H}}_{b}/\frac{w_{b}}{\alpha}) = -\alpha \mathbf{A}' \mathbf{R}^{-1} \mathbf{t} (\frac{\hat{\mathbf{n}}_{t}^{\mathrm{T}}}{d_{t}} - \frac{\hat{\mathbf{n}}_{b}^{\mathrm{T}}}{d_{b}}) \mathbf{A}^{-1}$$
(4.31)

which is about a fixed-value transformation (for fixed A,A',R,t) of $(\frac{\hat{\mathbf{n}}_t}{d_t} - \frac{\hat{\mathbf{n}}_b}{d_b})$. In other words, the difference of the two planes Π_t and Π_b could be captured by a matrix we refer to as the Biplanar Disparity Matrix D:

$$\mathbf{D} = \frac{\hat{\mathbf{H}}_{t}}{\frac{w_{t}}{\alpha}} - \frac{\hat{\mathbf{H}}_{b}}{\frac{w_{b}}{\alpha}} \tag{4.32}$$

Such a Biplanar disparity measure D encodes the difference between two planes: the plane that contains the peaks of the bumps, and the plane that contains the bottom of the bumps. It is invariant to global transformation of the wafer, but variant to relative transformation between the two planes, and is thus a measure we could use to decide if a wafer has bump heights that meet the specifications or not.

4.2.2 The Geometric meaning of the Biplanar Disparity matrix

Assume:

Fundermental matrix
$$\cong \begin{bmatrix} f_{11} & f_{12} & f_{13} \\ f_{21} & f_{22} & f_{23} \\ f_{31} & f_{32} & f_{33} \end{bmatrix}$$

The Fundamental matrix [34] about the peak plane and that of the bottom plane is the same, So:

$$\hat{\mathbf{F}}\hat{\mathbf{H}}_{t,9\times1} = \begin{bmatrix} (f_{11}R_{11} + f_{12}R_{21} + f_{13}R_{31}) + (f_{11}t_x + f_{12}t_y + f_{13}t_z)p_t \\ (f_{11}R_{12} + f_{12}R_{22} + f_{13}R_{32}) + (f_{11}t_x + f_{12}t_y + f_{13}t_z)q_t \\ (f_{11}R_{13} + f_{12}R_{23} + f_{13}R_{33}) + (f_{11}t_x + f_{12}t_y + f_{13}t_z)c_t \\ (f_{21}R_{11} + f_{22}R_{21} + f_{23}R_{31}) + (f_{21}t_x + f_{22}t_y + f_{23}t_z)p_t \\ (f_{21}R_{12} + f_{22}R_{22} + f_{23}R_{32}) + (f_{21}t_x + f_{22}t_y + f_{23}t_z)q_t \\ (f_{21}R_{13} + f_{22}R_{23} + f_{23}R_{33}) + (f_{21}t_x + f_{22}t_y + f_{23}t_z)c_t \\ (f_{31}R_{11} + f_{32}R_{21} + f_{33}R_{31}) + (f_{31}t_x + f_{32}t_y + f_{33}t_z)p_t \\ (f_{31}R_{12} + f_{32}R_{22} + f_{33}R_{32}) + (f_{31}t_x + f_{32}t_y + f_{33}t_z)q_t \\ (f_{31}R_{13} + f_{32}R_{23} + f_{33}R_{33}) + (f_{31}t_x + f_{32}t_y + f_{33}t_z)c_t \end{bmatrix}$$

$$\hat{\mathbf{F}}\hat{\mathbf{H}}_{b,9\times1} = \begin{bmatrix} (f_{11}R_{11} + f_{12}R_{21} + f_{13}R_{31}) + (f_{11}t_x + f_{12}t_y + f_{13}t_z)p_b \\ (f_{11}R_{12} + f_{12}R_{22} + f_{13}R_{32}) + (f_{11}t_x + f_{12}t_y + f_{13}t_z)q_b \\ (f_{11}R_{13} + f_{12}R_{23} + f_{13}R_{33}) + (f_{11}t_x + f_{12}t_y + f_{13}t_z)c_b \\ (f_{21}R_{11} + f_{22}R_{21} + f_{23}R_{31}) + (f_{21}t_x + f_{22}t_y + f_{23}t_z)p_b \\ (f_{21}R_{12} + f_{22}R_{22} + f_{23}R_{32}) + (f_{21}t_x + f_{22}t_y + f_{23}t_z)q_b \\ (f_{21}R_{13} + f_{22}R_{23} + f_{23}R_{33}) + (f_{21}t_x + f_{22}t_y + f_{23}t_z)c_b \\ (f_{31}R_{11} + f_{32}R_{21} + f_{33}R_{31}) + (f_{31}t_x + f_{32}t_y + f_{33}t_z)p_b \\ (f_{31}R_{12} + f_{32}R_{22} + f_{33}R_{32}) + (f_{31}t_x + f_{32}t_y + f_{33}t_z)q_b \\ (f_{31}R_{13} + f_{32}R_{23} + f_{33}R_{33}) + (f_{31}t_x + f_{32}t_y + f_{33}t_z)c_b \end{bmatrix}$$

assume the constant in the matrix into:

and so on so forth, the rewritten $\hat{\mathbf{F}}\hat{\mathbf{H}}_{t,9\times1}$ and $\hat{\mathbf{F}}\hat{\mathbf{H}}_{b,9\times1}$:

$$\hat{\mathbf{F}}\hat{\mathbf{H}}_{t,9\times1} = \begin{bmatrix} a_1 + \lambda_1 p_t & a_2 + \lambda_1 q_t & a_3 + \lambda_1 c_t \\ a_4 + \lambda_2 p_t & a_5 + \lambda_2 q_t & a_6 + \lambda_2 c_t \\ a_7 + \lambda_3 p_t & a_8 + \lambda_3 q_t & a_9 + \lambda_3 c_t \end{bmatrix}$$
(4.36)

$$\hat{\mathbf{F}}\hat{\mathbf{H}}_{b,9\times1} = \begin{bmatrix} a_1 + \lambda_1 p_b & a_2 + \lambda_1 q_b & a_3 + \lambda_1 c_b \\ a_4 + \lambda_2 p_b & a_5 + \lambda_2 q_b & a_6 + \lambda_2 c_b \\ a_7 + \lambda_3 p_b & a_8 + \lambda_3 q_b & a_9 + \lambda_3 c_b \end{bmatrix}$$
(4.37)

Where a1a2 a3 a4 a5 a6 a7 a8 a9 respectively represent the 9 constant in the matrix which are only related to the fundamental matrix [33] and the rotation matrix between two cameras[11]. If the cameras are fixed, then the intrinsic parameter A and translation between the two cameras are fixed, so B can be rewritten into:

$$B = (\mathbf{A}^{-1})^{\mathrm{T}} [\mathbf{t}]_{\times} \mathbf{A}^{-1} = \begin{bmatrix} b_1 & b_2 & b_3 \\ b_4 & b_5 & b_6 \\ b_7 & b_8 & b_9 \end{bmatrix}$$
(4.38)

then, $\frac{w_t}{\alpha}$ can be written with the fundamental matrix, the rotation and translation between two cameras and the intrinsic parameters[28] of the two cameras, as well as the 3D peak plane parameter. The $\frac{w_b}{\alpha}$ follows the same way.

$$\frac{w_{t}}{\alpha} = \sum_{i=1}^{9} a_{i}b_{i} + (b_{1}\lambda_{1} + b_{4}\lambda_{2} + b_{1}\lambda_{3})p_{t} + (b_{2}\lambda_{1} + b_{5}\lambda_{2} + b_{8}\lambda_{3})q_{t} + (b_{3}\lambda_{1} + b_{6}\lambda_{2} + b_{9}\lambda_{3})c_{t}$$

$$(4.39)$$

$$\frac{w_{\rm b}}{\alpha} = \sum_{i=1}^{9} a_i b_i + (b_1 \lambda_1 + b_4 \lambda_2 + b_1 \lambda_3) p_b + (b_2 \lambda_1 + b_5 \lambda_2 + b_8 \lambda_3) q_b + (b_3 \lambda_1 + b_6 \lambda_2 + b_9 \lambda_3) c_b$$
(4.40)

And can be simply expressed as:

$$\frac{w_t}{\alpha} = \phi + \kappa_1 p_t + \kappa_2 q_t + \kappa_3 c_t \tag{4.41}$$

$$\frac{w_b}{\alpha} = \phi + \kappa_1 p_b + \kappa_2 q_b + \kappa_3 c_b \tag{4.42}$$

Where ϕ , κ_1 , κ_2 , κ_3 are constant, which are only related to the parameters of rotation and translation between the two cameras, the camera intrinsic parameter [32]. And p, q, c is the 3D plane parameter in this article. Then

the Biplanar Disparity matrix is:

$$\mathbf{D} = \frac{\hat{\mathbf{H}}_{t}}{\frac{\mathbf{W}_{t}}{\alpha}} - \frac{\hat{\mathbf{H}}_{b}}{\frac{\mathbf{W}_{b}}{\alpha}} \\ = \begin{bmatrix} \frac{R_{11} + t_{x}p_{t}}{\phi + \kappa_{1}p_{t} + \kappa_{2}q_{t} + \kappa_{3}c_{t}} & \frac{R_{12} + t_{x}q_{t}}{\phi + \kappa_{1}p_{t} + \kappa_{2}q_{t} + \kappa_{3}c_{t}} & \frac{R_{13} + t_{x}c_{t}}{\phi + \kappa_{1}p_{t} + \kappa_{2}q_{t} + \kappa_{3}c_{t}} \\ \frac{R_{21} + t_{y}p_{t}}{\phi + \kappa_{1}p_{t} + \kappa_{2}q_{t} + \kappa_{3}c_{t}} & \frac{R_{22} + t_{y}q_{t}}{\phi + \kappa_{1}p_{t} + \kappa_{2}q_{t} + \kappa_{3}c_{t}} & \frac{R_{23} + t_{y}c_{t}}{\phi + \kappa_{1}p_{t} + \kappa_{2}q_{t} + \kappa_{3}c_{t}} \\ \frac{R_{31} + t_{x}p_{t}}{\phi + \kappa_{1}p_{t} + \kappa_{2}q_{t} + \kappa_{3}c_{t}} & \frac{R_{32} + t_{x}q_{t}}{\phi + \kappa_{1}p_{t} + \kappa_{2}q_{t} + \kappa_{3}c_{t}} & \frac{R_{33} + t_{x}c_{t}}{\phi + \kappa_{1}p_{t} + \kappa_{2}q_{t} + \kappa_{3}c_{t}} \end{bmatrix}$$

$$- \begin{bmatrix} \frac{R_{11} + t_{x}p_{b}}{\phi + \kappa_{1}p_{b} + \kappa_{2}q_{b} + \kappa_{3}c_{b}} & \frac{R_{12} + t_{x}q_{b}}{\phi + \kappa_{1}p_{b} + \kappa_{2}q_{b} + \kappa_{3}c_{b}} & \frac{R_{13} + t_{x}c_{b}}{\phi + \kappa_{1}p_{b} + \kappa_{2}q_{b} + \kappa_{3}c_{b}} \\ \frac{R_{21} + t_{y}p_{b}}{\phi + \kappa_{1}p_{b} + \kappa_{2}q_{b} + \kappa_{3}c_{b}} & \frac{R_{22} + t_{y}q_{b}}{\phi + \kappa_{1}p_{b} + \kappa_{2}q_{b} + \kappa_{3}c_{b}} & \frac{R_{33} + t_{y}c_{b}}{\phi + \kappa_{1}p_{b} + \kappa_{2}q_{b} + \kappa_{3}c_{b}} \\ \frac{R_{31} + t_{y}p_{b}}{\phi + \kappa_{1}p_{b} + \kappa_{2}q_{b} + \kappa_{3}c_{b}} & \frac{R_{32} + t_{x}q_{b}}{\phi + \kappa_{1}p_{b} + \kappa_{2}q_{b} + \kappa_{3}c_{b}} & \frac{R_{33} + t_{y}c_{b}}{\phi + \kappa_{1}p_{b} + \kappa_{2}q_{b} + \kappa_{3}c_{b}} \end{bmatrix}$$

If the peak plane of the bumps are parallel to that of the bottom plane, which means:

$$\frac{p_t}{p_b} = \frac{q_t}{q_b} = \frac{c_t}{c_b} = \tau \tag{4.44}$$

then the bump height will be the distance of the two planes. Represent $\kappa_1 p_t + \kappa_2 q_t + \kappa_3 c_t$ to be ε . Then the Biplanar Disparity matrix changes into:

$$D = \begin{bmatrix} \frac{(\phi t_x p_t - \varepsilon R_{11}) - \tau(\phi t_x p_t - \varepsilon R_{11})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} & \frac{(\phi t_x q_t - \varepsilon R_{12}) - \tau(\phi t_x q_t - \varepsilon R_{12})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} & \frac{(\phi t_x c_t - \varepsilon R_{13}) - \tau(\phi t_x c_t - \varepsilon R_{13})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} \\ \frac{(\phi t_x p_t - \varepsilon R_{21}) - \tau(\phi t_x p_t - \varepsilon R_{21})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} & \frac{(\phi t_x q_t - \varepsilon R_{22}) - \tau(\phi t_x q_t - \varepsilon R_{22})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} & \frac{(\phi t_x c_t - \varepsilon R_{23}) - \tau(\phi t_x c_t - \varepsilon R_{23})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} \\ \frac{(\phi t_x p_t - \varepsilon R_{31}) - \tau(\phi t_x p_t - \varepsilon R_{31})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} & \frac{(\phi t_x q_t - \varepsilon R_{32}) - \tau(\phi t_x q_t - \varepsilon R_{33})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} \\ \frac{(\phi t_x p_t - \varepsilon R_{11}) - \tau(\phi t_x p_t - \varepsilon R_{22})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} & \frac{(\phi t_x q_t - \varepsilon R_{23}) - \tau(\phi t_x q_t - \varepsilon R_{23})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} \\ \frac{(\phi t_x p_t - \varepsilon R_{31}) - \tau(\phi t_x c_t - \varepsilon R_{33})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} & \frac{(\phi t_x c_t - \varepsilon R_{33}) - \tau(\phi t_x c_t - \varepsilon R_{33})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} \\ \frac{(\phi t_x p_t - \varepsilon R_{11}) - \tau(\phi t_x q_t - \varepsilon R_{12})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} & \frac{(\phi t_x q_t - \varepsilon R_{33}) - \tau(\phi t_x c_t - \varepsilon R_{33})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} \\ \frac{(\phi t_x p_t - \varepsilon R_{11}) - \tau(\phi t_x q_t - \varepsilon R_{12})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} & \frac{(\phi t_x q_t - \varepsilon R_{33}) - \tau(\phi t_x c_t - \varepsilon R_{33})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} \\ \frac{(\phi t_x p_t - \varepsilon R_{33}) - \tau(\phi t_x c_t - \varepsilon R_{33})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} \\ \frac{(\phi t_x p_t - \varepsilon R_{33}) - \tau(\phi t_x c_t - \varepsilon R_{33})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} \\ \frac{(\phi t_x p_t - \varepsilon R_{33}) - \tau(\phi t_x c_t - \varepsilon R_{33})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} \\ \frac{(\phi t_x p_t - \varepsilon R_{33}) - \tau(\phi t_x c_t - \varepsilon R_{33})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} \\ \frac{(\phi t_x p_t - \varepsilon R_{33}) - \tau(\phi t_x c_t - \varepsilon R_{33})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} \\ \frac{(\phi t_x p_t - \varepsilon R_{33}) - \tau(\phi t_x c_t - \varepsilon R_{33})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} \\ \frac{(\phi t_x p_t - \varepsilon R_{33}) - \tau(\phi t_x c_t - \varepsilon R_{33})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} \\ \frac{(\phi t_x p_t - \varepsilon R_{33}) - \tau(\phi t_x c_t - \varepsilon R_{33})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} \\ \frac{(\phi t_x p_t - \varepsilon R_{33}) - \tau(\phi t_x c_t - \varepsilon R_{33})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} \\ \frac{(\phi t_x p_t - \varepsilon R_{33}) - \tau(\phi t_x c_t - \varepsilon R_{33})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} \\ \frac{(\phi t_x p_t - \varepsilon R_{33}) - \tau(\phi t_x c_t - \varepsilon R_{33})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} \\ \frac{(\phi t_x p_t - \varepsilon R_{33}) - \tau(\phi t_x c_t - \varepsilon R_{33})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} \\ \frac{(\phi t_x p_t - \varepsilon R_{33$$

If the bump height has some global changes, which means, all the parameters p_t, q_t, c_t unchanged, only τ has some changes which are inversely proportional to the bump height.

From the D matrix, we can see that the changes of the bump height have some global influence on every 9 value with respect to the changes of , the norm of the Biplanar Disparity matrix will have some close-to-linear translation[60] against the bump height, which is the same as the analysis by using the

synthetic data. (Fig. 4.1). To simple the question, we induce the circumstance

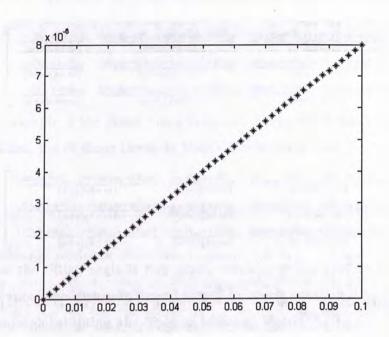


Figure 4.1: theoretical analysis on the change in norm of Biplanar Disparity with respect to bump height change from 20 to 100 microns[29]

in the real inspection experiment. The reference coordinate is the camera coordinate, the feed-in system working plane is on the parallel coordinate, so in the expression of the Biplanar Disparity Matrix D, pt and qt are 0. Then the new expression of D will be:

$$D = \begin{bmatrix} \frac{-\varepsilon R_{11} + \tau \varepsilon R_{11}}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} & \frac{-\varepsilon R_{12} + \tau \varepsilon R_{12}}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} & \frac{(\phi t_x c_t - \varepsilon R_{13}) - \tau (\phi t_x c_t - \varepsilon R_{13})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} \\ \frac{-\varepsilon R_{21} + \tau \varepsilon R_{21}}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} & \frac{-\varepsilon R_{22} + \tau \varepsilon R_{22}}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} & \frac{(\phi t_x c_t - \varepsilon R_{23}) - \tau (\phi t_x c_t - \varepsilon R_{23})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} \\ \frac{-\varepsilon R_{31} + \tau \varepsilon R_{31}}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} & \frac{-\varepsilon R_{32} + \tau \varepsilon R_{32}}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} & \frac{(\phi t_x c_t - \varepsilon R_{33}) - \tau (\phi t_x c_t - \varepsilon R_{33})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} \end{bmatrix}$$

$$(4.46)$$

Then in this way, the influence on the 3rd column of the matrix is the biggest with respect to the change of bump height.

If the bump height has no changes, only the peak plane and the bottom plane of the bumps both have some global tilting, which means, only change, but is maintain, the 9 value in the Biplanar Disparity matrix will have different changes.

1) For example, if both the two planes have global tilting around x axis, which means pt is equal to 0. Then the Biplanar Disparity Matrix D changes into:

$$D = \begin{bmatrix} \frac{-\varepsilon R_{11} + \tau \varepsilon R_{11}}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} & \frac{(\phi t_x q_t - \varepsilon R_{12}) - \tau(\phi t_x q_t - \varepsilon R_{12})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} & \frac{(\phi t_x c_t - \varepsilon R_{13}) - \tau(\phi t_x c_t - \varepsilon R_{13})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} \\ \frac{-\varepsilon R_{21} + \tau \varepsilon R_{21}}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} & \frac{(\phi t_x q_t - \varepsilon R_{22}) - \tau(\phi t_x q_t - \varepsilon R_{22})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} & \frac{(\phi t_x c_t - \varepsilon R_{23}) - \tau(\phi t_x c_t - \varepsilon R_{23})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} \\ \frac{-\varepsilon R_{31} + \tau \varepsilon R_{31}}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} & \frac{(\phi t_x q_t - \varepsilon R_{32}) - \tau(\phi t_x q_t - \varepsilon R_{22})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} & \frac{(\phi t_x c_t - \varepsilon R_{33}) - \tau(\phi t_x c_t - \varepsilon R_{33})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} \end{bmatrix}$$

$$(4.47)$$

2) For example, if the global tilting is around y axis, which means qt is equal to 0. Then, the Biplanar Disparity Matrix D will change into:

$$D = \begin{bmatrix} \frac{(\phi t_x p_t - \varepsilon R_{11}) - \tau(\phi t_x p_t - \varepsilon R_{11})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} & \frac{-\varepsilon R_{12} + \tau \varepsilon R_{12}}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} & \frac{(\phi t_x c_t - \varepsilon R_{13}) - \tau(\phi t_x c_t - \varepsilon R_{13})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} \\ \frac{(\phi t_x p_t - \varepsilon R_{21}) - \tau(\phi t_x p_t - \varepsilon R_{21})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} & \frac{-\varepsilon R_{22} + \tau \varepsilon R_{22}}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} & \frac{(\phi t_x c_t - \varepsilon R_{23}) - \tau(\phi t_x c_t - \varepsilon R_{23})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} \\ \frac{(\phi t_x p_t - \varepsilon R_{31}) - \tau(\phi t_x p_t - \varepsilon R_{31})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} & \frac{-\varepsilon R_{32} + \tau \varepsilon R_{32}}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} & \frac{(\phi t_x c_t - \varepsilon R_{33}) - \tau(\phi t_x c_t - \varepsilon R_{33})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} \end{bmatrix}$$

$$(4.48)$$

Because the tilting angle is very small, which is within 1/60 as said in the specification, the change in pt and qt will be nearly equal. As in example 1), after some global tilting, the Biplanar Disparity Matrix will be:

$$D = \begin{bmatrix} \frac{-\varepsilon R_{11} + \tau \varepsilon R_{11}}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} & \frac{(\phi t_x \Delta q_t - \tau \phi t_x \Delta q_t) - (\varepsilon R_{12} + \tau \varepsilon R_{12})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} & \frac{(\phi t_x \Delta c_t - \tau \phi t_x \Delta c_t) - (\varepsilon R_{13} + \tau \varepsilon R_{13})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} \\ \frac{-\varepsilon R_{21} + \tau \varepsilon R_{21}}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} & \frac{(\phi t_x \Delta q_t - \tau \phi t_x \Delta q_t) - (\varepsilon R_{22} + \tau \varepsilon R_{22})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} & \frac{(\phi t_x \Delta c_t - \tau \phi t_x \Delta c_t) - (\varepsilon R_{23} + \tau \varepsilon R_{23})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} \\ \frac{-\varepsilon R_{31} + \tau \varepsilon R_{31}}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} & \frac{(\phi t_x \Delta q_t - \tau \phi t_x \Delta q_t) - (\varepsilon R_{32} + \tau \varepsilon R_{32})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} & \frac{(\phi t_x \Delta c_t - \tau \phi t_x \Delta c_t) - (\varepsilon R_{33} + \tau \varepsilon R_{33})}{(\phi + \varepsilon)(\phi + \tau \varepsilon)} \end{bmatrix}$$

$$(4.49)$$

$$\|D'\| = \begin{cases} \left(\frac{-\varepsilon R_{11} + \tau \varepsilon R_{11}}{(\phi + \varepsilon)(\phi + \tau \varepsilon)}\right)^2 + \left(\frac{-\varepsilon R_{21} + \tau \varepsilon R_{21}}{(\phi + \varepsilon)(\phi + \tau \varepsilon)}\right)^2 + \left(\frac{\varepsilon R_{32} - \tau \varepsilon R_{32}}{(\phi + \varepsilon)(\phi + \tau \varepsilon)}\right)^2 \\ + \left(\frac{\varepsilon R_{12} + \tau \varepsilon R_{12}}{(\phi + \varepsilon)(\phi + \tau \varepsilon)}\right)^2 + \left(\frac{\varepsilon R_{22} - \tau \varepsilon R_{22}}{(\phi + \varepsilon)(\phi + \tau \varepsilon)}\right)^2 + \left(\frac{\varepsilon R_{32} - \tau \varepsilon R_{32}}{(\phi + \varepsilon)(\phi + \tau \varepsilon)}\right)^2 \\ + \left(\frac{\varepsilon R_{13} - \tau \varepsilon R_{13}}{(\phi + \varepsilon)(\phi + \tau \varepsilon)}\right)^2 + \left(\frac{\varepsilon R_{23} - \tau \varepsilon R_{23}}{(\phi + \varepsilon)(\phi + \tau \varepsilon)}\right)^2 + \left(\frac{\varepsilon R_{33} - \tau \varepsilon R_{33}}{(\phi + \varepsilon)(\phi + \tau \varepsilon)}\right)^2 \\ + 3\left(\frac{\phi t_x q_t - \tau \phi t_x q_t}{(\phi + \varepsilon)(\phi + \tau \varepsilon)}\right)^2 \Delta^2 + 3\left(\frac{\phi t_x c_t - \tau \phi t_x c_t}{(\phi + \varepsilon)(\phi + \tau \varepsilon)}\right)^2 \Delta^2 \\ - 2\Delta\phi t_x q_t \varepsilon (1 - \tau^2) \left(\frac{1}{(\phi + \varepsilon)(\phi + \tau \varepsilon)}\right) \left(\frac{1}{(\phi + \varepsilon)(\phi + \tau \varepsilon)}\right) R_{12} \\ - 2\Delta\phi t_x q_t \varepsilon (1 - \tau^2) \left(\frac{1}{(\phi + \varepsilon)(\phi + \tau \varepsilon)}\right) \left(\frac{1}{(\phi + \varepsilon)(\phi + \tau \varepsilon)}\right) R_{32} \\ - 2\Delta\phi t_x c_t \varepsilon (1 - \tau^2) \left(\frac{1}{(\phi + \varepsilon)(\phi + \tau \varepsilon)}\right) \left(\frac{1}{(\phi + \varepsilon)(\phi + \tau \varepsilon)}\right) R_{13} \\ - 2\Delta\phi t_x c_t \varepsilon (1 - \tau^2) \left(\frac{1}{(\phi + \varepsilon)(\phi + \tau \varepsilon)}\right) \left(\frac{1}{(\phi + \varepsilon)(\phi + \tau \varepsilon)}\right) R_{23} \\ \sqrt{-2\Delta\phi t_x c_t \varepsilon (1 - \tau^2)} \left(\frac{1}{(\phi + \varepsilon)(\phi + \tau \varepsilon)}\right) \left(\frac{1}{(\phi + \varepsilon)(\phi + \tau \varepsilon)}\right) R_{33} \end{cases}$$

Here $\Delta=1+\delta$, and δ is much close to zero. So, the difference between $\|D\|$ and $\|D\|$ will be little. That means, when there is only global tilting, but no change in bump height, the Norm of the Biplanar Disparity Matrix will have little change, which will be testified in the synthetic data experiment.

4.3 RANSAC-Random Sample Consensus

The RANSAC[70][68] procedure is opposite to that of conventional smoothing techniques: Rather than using as much of the data as possible to obtain an initial solution and then attempting to eliminate the invalid data points, RANSAC uses as small an initial data set as feasible and enlarges this set with consistent data when possible. For example, given the task of fitting an arc of a circle to a set of two-dimensional points, the RANSAC approach[67] would be to select a set of three points (since three points are required to determine a circle)[71], compute the center and radius of the implied circle, and count the number of points that are close enough to that circle to suggest compatibility with it (i.e., their deviations are small enough to be measurement errors). If there are enough compatible points, RANSAC would employ a smoothing technique such as least squares, to compute an improved estimate for the parameters of the circle now that a set of mutually consistent points has been identified.

4.3.1 finding Homography matrix by using RANSAC[72][35]

Once I use RANSAC, I get a Homography, then use that Homography and feature positions in one view to predict feature positions in the second view, and check the distance between predicted positions and observed positions in the second view. Use those distances to decide if any particular feature correspondence pair is outlier or not. In this thesis, to calculate a Homography matrix

there needs at least 4 pairs of correspondences. By using RANSAC method, the system can recognize the outliers by itself.

4.3.2 finding Fundamental matrix by using RANSAC[73][34]

The fundamental matrix F relates points in two images. If x is a point in one image and x' a point in another image, then $(x')^TFx = 0$. To compute F completely automatically we begin by using a corner detector to find interest points in an image. Putative matches of the feature points in both images are computed by using a correlation measure[38] for points in one image with the features in the other image. Only features within a small window are considered to limit computation time. Mutually best matches are retained.

RANSAC is used to robustly determine F from these putative matches. A minimal sample is selected from the putative matches from which a tentative F is calculated. The process is iterated until a sufficient number of samples have been taken. The F which fits the matches that becomes the input for the next step.

Once an initial F has been computed, more matches can be found by searching along epipolar lines[3]. A non-linear minimization is used to fit an F to a large number of points.

4.4 Harris Corner detection

In last subsection of finding the fundamental matrix, we mentioned about finding the initial putative correspondence points by using Corner Detection[21][75]. In this thesis, we use Harris Corner Detection to find these feature points. Points whose intensity gradient is bigger than a threshold are pointed to be corners. To each choice of feature point on the substrate, we apply corner detection[74] first, then hand-pick a small area, finally auto-search the nearest corner[74] in this area to be the exact point. All the feature points on the substrate are

extracted by this way.

Chapter 5

Simulation and experimental results

5.1 Simulation experiments

5.1.1 Preliminary experiments

For each bump height under a particular setting on the errors and uncertainties, we allow the errors and uncertainties to fluctuate randomly within some predefined range of values for 40 times ranging from 20 microns to 100 microns, and the average of the norm and direction fluctuation of the Disparity Matrix D are plotted against the bump height.

Without any error or uncertainty considered, i.e., in the most ideal case, norm of the Disparity Matrix D has the following behavior. There is also zero fluctuation in the direction of D (when D is viewed as 9-D vector).

In the following series of simulation images, the x axis means the changing of bump height, and the y axis means the norm of the corresponding Norm of Disparity Matrix D. The following figures present how the norm fluctuation of the D matrix behave with changing bump height, under different combinations of errors and uncertainties.

If the golden wafer is 60 microns in height, then \pm 20 microns is the allowed

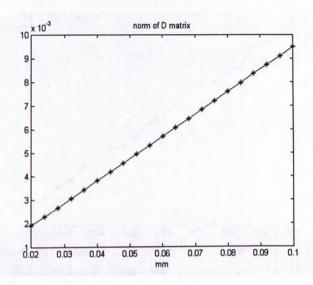


Figure 5.1: Norm of D in the absence of any error or uncertainty

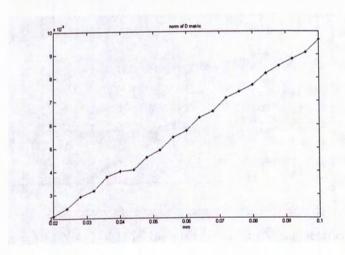


Figure 5.2: Norm of D in the presence of image resolution error

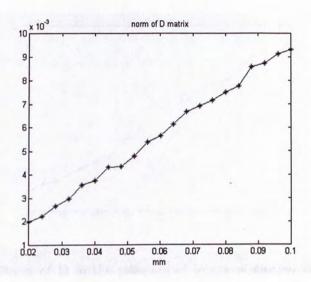


Figure 5.3: Norm of D in the presence of image resolution error and bump position variation uncertainty

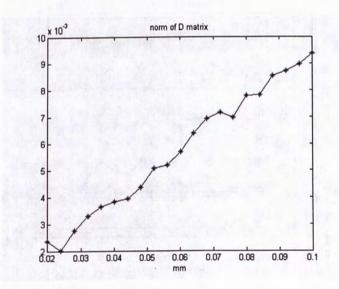


Figure 5.4: Norm of D in the presence of image resolution error, brightest point determination uncertainty, and bump position variation uncertainty

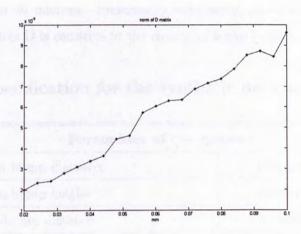


Figure 5.5: Norm of D in the presence of image resolution error, bump position variation uncertainty and wafer-plane transformation and wafer-tilting transformation

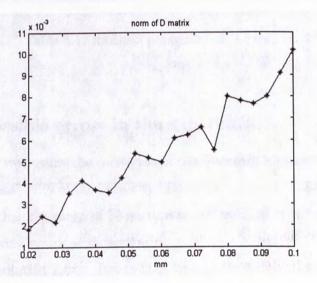


Figure 5.6: Norm of D in the presence of image resolution error, brightest point determination uncertainty, bump position variation uncertainty, and wafer-tilting transformation.

error. So, we only need to detect those bumps who are lower than 40 microns or bigger than 80 microns. Preliminary experiment shows that the Norm of Disparity matrix D is sensitive to the change of bump heights.

5.1.2 Specification for the synthetic data system

Parameters of the system				
Golden bump diameter	120micron			
Golden bump height	60micron			
Bump-bump distance	70 micron			
Numbers of bumps per-inspection	20			
Distance between wafer to camera frame	91mm			
Angle between two light sources	15(degree)			
Resolution of CCD	500*500/ 300 circles			
x, y resolution	30 micron			
z resolution	20 micron			

Table 5.1: Intrinsic parameter of the camera

5.1.3 Possible errors in the experiment

In this thesis we outline the derivation of the Disparity Matrix D. We also concern on the sensitivity of the Disparity Matrix with respect to the average bump height which has the range of 20 microns to 100 microns in our experiment under the following errors or uncertainties:

- 1) Image Resolution Error: Image is of limited resolution, thus image projection will be truncated to the position of the nearest pixel in the image. In our experiments we allow such error to be in the range \pm 1 pixel in each image dimension of the true position.
- 2) Brightest Point Determination Uncertainty: Under our illumination system the peaks of the bumps would appear as bright points in the image. However,

due to image sensing saturation, feature extraction, or other factors, there could be uncertainty in determining the position of such brightest points in the image. In our experiments, we assume such uncertainty to be in the range \pm 1 pixels in each image dimension of the true position which mostly happens in the real inspection process.

- 3) Bump Position Variation Uncertainty: The bump array on a wafer could have a few variations in the bump positions on the wafer plane due to manufacturing and other uncontrollable factors. We allow each bump to vary by \pm 5mm of the nominal position.
- 4) Wafer Global Transformation: Ideally, the disparity matrix should be invariant with rigid transformation of the wafer. To ease our experiments, we consider the global transformation in two groups:
- (a) Wafer-plane transformation: there is only translation and rotation on the wafer plane (Range of such transformation that is considered in our experiments:0 to 500 microns in translation, and 0 to 2 degree in rotation);
- (b) Wafer-tilting Transformation: there could be tilting of the wafer about its own plane (Range of such transformation that is considered in our experiments: the biggest tilt height is 1mm). Such transformation could cause the brightest points of the bumps (due to specular reflections of the bump surface under parallel or point source of illumination) to have their positions shifted on the bump surfaces, and could pose a bigger challenge than the above wafer-plane transformation.

5.2 Real images experiments

5.2.1 Experiment instrument

The real image data experiment should meet the requirement below, and set up as shown in Fig. 5.7 and Fig. 5.8

Image Distance of Camera	131.7105mm
Distance of bump bottom plane b from camera	91mm
Image Resolution	768*576 in pixel
Angle between the two parallel sources of Illumination	15 degree

Table 5.2: Extrinsic Parameter of the cameras and CCD resolution



Figure 5.7: for center imaging part

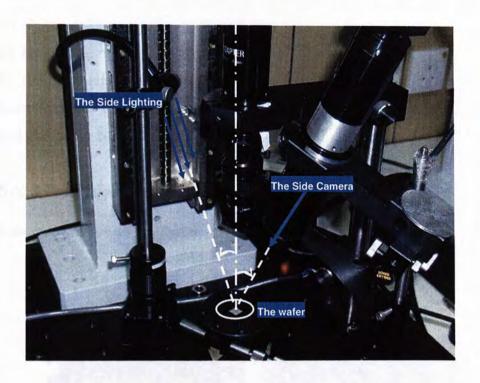


Figure 5.8: for side imaging part

5.2.2 The Inspection Procedure

- 1. Set up the imaging system and do the calibration. [14]
- 2. For reference wafer:
- (a) Point Correspondence Acquisition: Get image point correspondences $\{(\mathbf{p}_i^t, \mathbf{p'}_i^t) : i = 1, 2, \cdots\}$ over the bump peaks, and $\{(\mathbf{p}_j^b, \mathbf{p'}_j^b) : j = 1, 2, \cdots\}$ over the bump bottoms.
- (b) Homography Estimation: Estimate the Homographies \mathbf{H}_{t} and \mathbf{H}_{b} using Equations 4.9 and 4.10, from the two sets of point correspondences[39].
- (c) Standard D Acquisition: Use Equation 4.23 to measure $\frac{w_t}{\alpha}$ and $\frac{w_b}{\alpha}$ for $\hat{\mathbf{H}}_t$ and $\hat{\mathbf{H}}_b$ respectively. Use Equation 4.32 to come up with a standard value D for the reference wafer. 3. For incoming wafer:
- (a) Point Correspondence Acquisition: Get image point correspondences $\{(\mathbf{p}_i^t, \mathbf{p'}_i^t) : i = 1, 2, \cdots\}$ over the bump peaks, and $\{(\mathbf{p}_j^b, \mathbf{p'}_j^b) : j = 1, 2, \cdots\}$ over the bump bottoms.
- (b) Homography Estimation: Estimate the Homographies \mathbf{H}_{t} and \mathbf{H}_{b} using Equations 4.9 and 4.10, from the two sets of point correspondences.

(c) D Acquisition and Comparison: Use Equation 4.23 to measure $\frac{w_b}{\alpha}$ and $\frac{w_b}{\alpha}$ for $\hat{\mathbf{H}}_t$ and $\hat{\mathbf{H}}_b$ respectively. Use Equation 4.32 to come up with a measure of D for the incoming wafer. Compare this D with the standard D of the reference wafer to decide if for this incoming wafer the difference of the bump peak plane and the bump bottom plane is within acceptable range.

5.2.3 Images grabbed under above system

Images under the center camera and center lighting, as the red camera in Fig.1.1

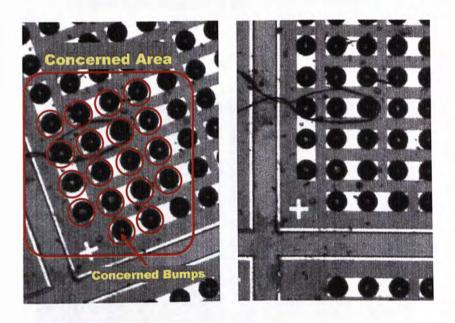


Figure 5.9: Real image group 1-center images:1,2

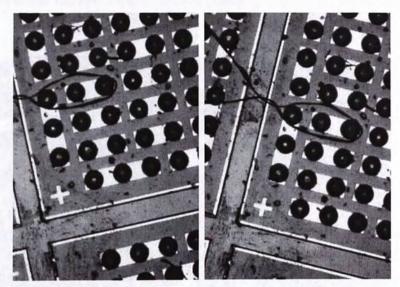


Figure 5.10: Real image group 1-center images:3,4

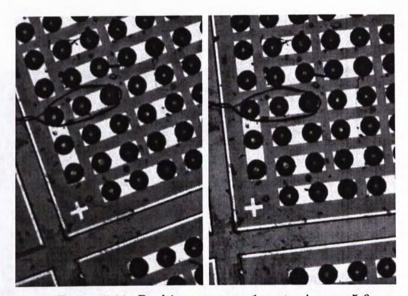


Figure 5.11: Real image group 1-center images:5,6

Images under the side camera and side lighting, as the blue camera in Fig.1.1

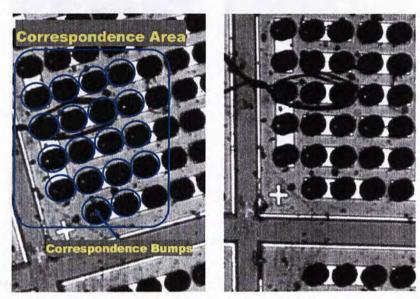


Figure 5.12: Real image group 1-side images:1,2

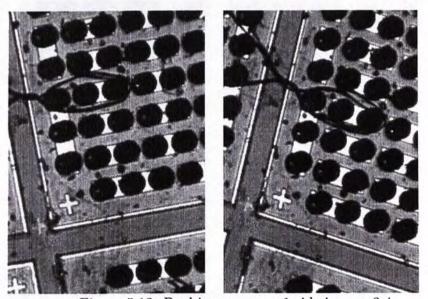


Figure 5.13: Real image group 1-side images:3,4

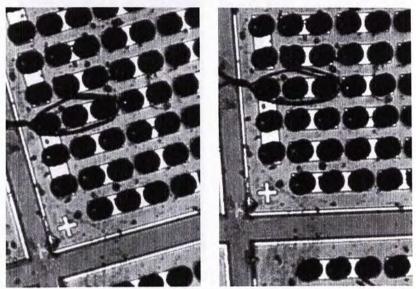


Figure 5.14: Real image group 1-side images: 5,6

Images under the center camera and center lighting, as the red camera in Fig.1.1

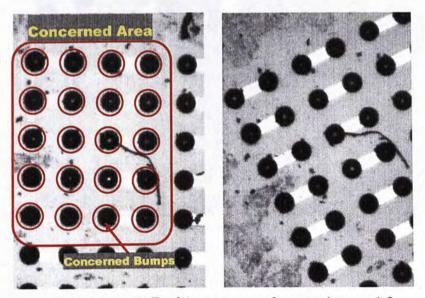


Figure 5.15: Real image group 2-center images:1,2

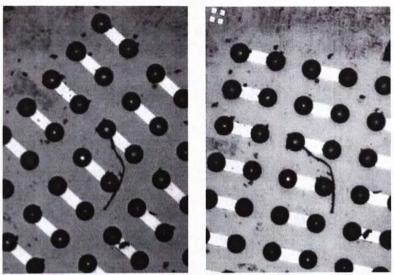


Figure 5.16: Real image group 2-center images:3,4

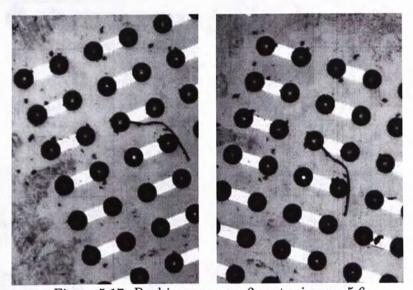


Figure 5.17: Real image group 2-center images:5,6

Images under the side camera and side lighting, as the blue camera in Fig.1.1 $\,$

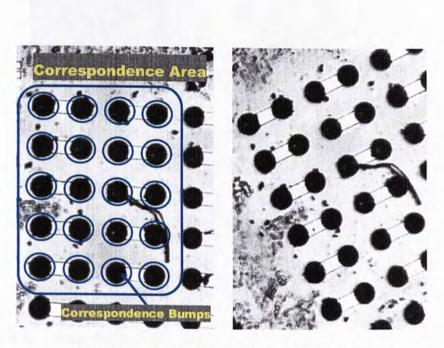


Figure 5.18: Real image group 2-side images:1,2

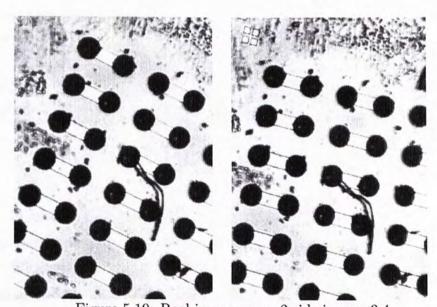


Figure 5.19: Real image group 2-side images:3,4

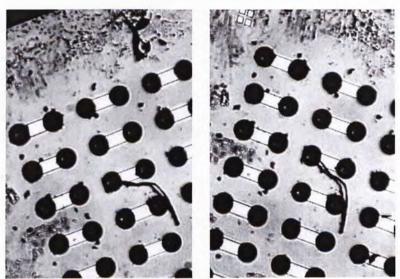


Figure 5.20: Real image group 2-side images: 5,6

5.2.4 Experimental Results

The experiment results are shown in the following table. The numbers here represent the Norm of the Biplanar Disparity matrix from different images named from figure 1 to figure 6 as shown above. There are two different wafers divided into group 1 and group 2.

	Figure 1	Figure 2	Figure 3	Figure 4	Figure 5	Figure 6	mean
group 1	0.175251	0.163423	0.183401	0.180018	0.176767	0.168152	0.1745
group 2	0.263179	0.266873	0.281037	0.264023	0.267192	0.275998	0.2681

Table 5.3: different norms with respect to different wafer

In geometric express of the above data is two separate curve which can be separated clearly into two different groups as the figure below.

Within each group, we find that the norm of Biplanar Disparity D. is nearly stable, the D's mean of group one is 0.1782, correspondingly, which is 0.2831

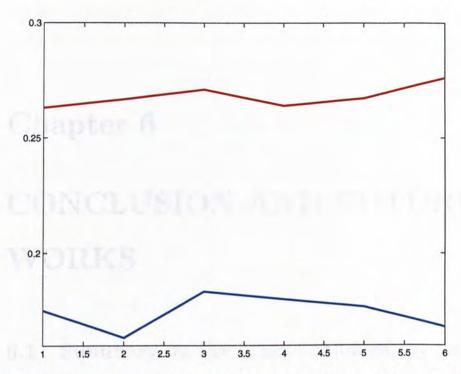


Figure 5.21: two groups of D. Norm value about two pieces of Die with different bump size

Fluctuation	Figure 1	Figure 2	Figure 3	Figure 4	Figure 5	Figure 6	1 and 2
within group1	0.43%	-6.35%	2.85%	5.10%	3.16%	-3.64%	53.61%
within group2	-1.82%	-0.44%	1.11%	-1.50%	-0.32%	2.96%	

Table 5.4: changing ratio of the norms' value within the two groups

of group two. However, across different groups, compared with the fluctuation in a single group, the difference is much bigger between different wafer, which is because the height of these two kinds of wafer is quite different, and the Biplanar Disparity D. just somewhat describes the distance between the peak plane and the bottom plane of the bumps.

Chapter 6

CONCLUSION AND FUTURE WORKS

6.1 Summary on the contribution of my work

From the above experimental results, both the synthetic experiment and the real data experiment show that the Biplanar Disparity matrix is sensitive enough to the change in bump height, even under a number of errors and uncertainties. These results can preliminary testify that the Biplanar Disparity matrix D. which includes the planar information of both of the peak plane and the bottom plane of the bumps, has a tight relationship with the relative position of these two planes. According to the real world false bump inspection, two cases happen mostly. One case is: when there is no change in the bottom plane, only some parallel change in the peak plane, which indicates only the bump height is changing, the results show the norm of Biplanar Disparity D. will change at the same time. The other case is: when there is no distance change between the two planes, but only global titling of the two planes, which indicates that the bump height is constant, this case often happens when the glue between the substrate and the wafer has some change, results show that the norm of Biplanar Disparity D. will have little change. Our expected inspection system can mostly meet the need of our objectives for this work. Our hard ware for this system only needs two sets of cameras and lenses, as well as a feed-in system which altogether cost lower than two thousand US dollars and is much cheaper than the previous ones. Our Biplanar Disparity matrix measurement is and aim-direct approach. To examine the speed of the inspection process, under matlab command window, and without counting the time of feature extraction and finding corresponding points, our method needs 1.85 seconds for the inspection process which includes SVD, Robust estimation, Homography estimation, D. matrix estimation, and comparison of specimen D. and reference D. for inspecting about 400 bumps. While, under the same circumstance, the method of sinusoid grating needs 9.58 seconds. In a word, the Biplanar Disparity measurement approach needs not explicit 3D reconstruction, which not only reduce the time but also simplify the whole procedure of inspection. It is a fast approach for the industrial world and can meet most of the objectives we are going to achieve. To conclude that it is a promising method for the future 3D inspection of wafer bumping.

6.2 Some Weakness of The Method

If there are some bumps on the wafer having not only one peak, for example if there are two peaks on a single bump, in both central and side image of this bump, on the top area, there will be two brightest points. Then, in the process of finding out the outliers by using robust estimation, in the peak area, only one peak will be selected. But, if the false peak is brighter than the real peak, some error answer will be produced.

6.3 Future Works and Further Development

6.3.1 About the synthetic experiment

Fig.5.6shows that the norm of Biplanar Disparity D. could be affected to a significant extent by the brightest point determination uncertainty, and to a less extent by the tilting of the wafer. Future work will address how these could be alleviated. However, we must point out that, if higher image resolution could be

adopted, the sensitivity of D to the above could be much reduced. The following figure shows that if the image resolution is increased to 1000*1000[49], the effect of the above to the norm of D is minimal.

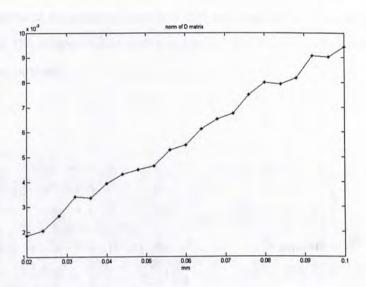


Figure 6.1: Norm of D in the presence of image resolution error, brightest point determination uncertainty, bump position variation uncertainty, and wafer-tilting transformation, under a higher image resolution 1000 X 1000

6.3.2 About the real image experiment

Firstly, in the side images of the bumps, there is some shadow which can not be eliminated now. This will lead the whole system not totally automatically to find the correspondence points on the substrate. Further research should be done to tackle this problem, for hand-picking the patterns on the substrate to be the correspondence points will more or less induce some errors, thus make the result rough. Secondly coming, not only the norm of the Biplanar Disparity matrix D can give the information of the two planes, but also the orientation of the D vector[60] (to reshape the Biplanar Disparity matrix into a 9 by 1 vector) can indicate more on the relative position of the two planes. For example, the

tilting angle, or around which axis the planes tilt[50]. Future work should focus on this area, and make the result afford more details about the specimen wafer. After that, the angle between the two cameras in both the synthetic experiment and the real image experiment should not only as rough as 15 degree, but need some scientific research on the optimized angle. Finally, the whole inspection system needs to be improved into a totally automatical one, including the feed in system, the camera calibration process, the grabbing images system, and the inspection process.

Bibliography

- Jorge L.C. Sanz and Ramesh Jain, "Advances in Machine Vision",
 Springer-Verlag New york Berlin Heigelberg London, pp. 165-171,1995.
- [2] Antonio Criminisi, C.J. van Rijsbergen, "Accurate visual metrology from single and multiple uncalibrated images", Springer-Verlag, pp. 25-34, 2001.
- [3] Gang Xu and Zhengyou Zhang, "Epipolar geometry in stereo, motion and object recognition", Kluwer academic publishers, Vol 6, pp. 79-117, 1996.
- [4] Stan Birchfield, "An introduction to projective geometry", Berlin: Springer-verlag, pp. 226-248, Oct. 1998.
- [5] Maybank, Stephen, "Theory of reconstruction from image motion", Berlin
 : Springer-verlag; Hong Kong: Springer-Verlag, c1993, pp.174-180.
- [6] Boyer, Kim LaVerne, "3-D robot vision: acquisition and interpretation", Ann Arbor, Mich.: University Microfilms International, 1988.
- [7] Siwook Nam, Hanjoo Kim, Jaihie Kim, Nicolai Petkov, Michel A. West-enberg, "Trajector Estimation Based on Globally Consistent Homography", Proceedings on Computer analysis of images and patterns: 10th international conference, CAIP 2003, Groningen, The Netherlands, pp.214-222.
- [8] Faugeras, O. and Mourrain, B., "About the correspondences of points between N images", Proceedings of IEEE, pp.37 44, 24 June 1995.
- [9] Huttenlocher, D.P. and Lorigo, L.M., "Recognizing three-dimensional objects by comparing two-dimensional images", IEEE Computer Society

- Conference on Vision and Pattern Recognition, 1996. Proceedings CVPR, pp.878 884, June 1996.
- [10] Werner, T. and Pajdla, T., "Cheirality in epipolar geometry", IEEE International Conference on Computer Vision, Proceedings of ICCV, Vol. 1, pp.548 553, July 2001.
- [11] Ding Yuan and Ronald Chung, "Correspondence-free stereo vision for the case of arbitrarily-positioned cameras", ICRA, pp. 1688-1693, 2003.
- [12] Frahm, J.-M. and Koch, R., "Camera calibration with known rotation", IEEE Proceedings on Computer Vision, Vol.2, pp.1418 - 1425,Oct. 2003.
- [13] Heikkila, J. and Silven, O., "A four-step camera calibration procedure with implicit image correction", *IEEE Proceedings on CVPR*, pp.1106 - 1112, June 1997.
- [14] Zhengyou Zhang, "Camera Calibration with One-Dimensional Objects", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 26,pp.892 - 899, July 2004.
- [15] Hanqi Zhuang and Wen-Chiang Wu, "Camera calibration with a nearparallel (ill-conditioned) calibration board configuration", *IEEE Transac*tions on Robotics and Automation, Vol. 12, Issue. 6, pp. 918 - 921, Dec. 1996.
- [16] Jong-Soo Lee and Yu-Ho Jeong, "CCD camera calibrations and projection error analysis", Science and Technology KORUS 2000 Proceedings on the 4th Korea-Russia International Symposium, Vol. 2,pp. 50 - 55, July 2000.
- [17] Reza Asgari, "Semiconductor backend flip chip processing, inspection requirements and challenges", Electronics Manufacturing Technology Symposium, pp. 18-22, June 2002.
- [18] Maur, F., "X-ray inspection for electronic packaging latest developments", Electronic Packaging Technology Proceedings 2003 ICEPT2003, pp. 235 -239, Oct. 2003.

- [19] Adriance, J.H.; Whitmore, M.A.; Schake, J.D., "Bumping of silicon wafers by stencil printing", Electronics Manufacturing Technology Symposium, Twenty-Fourth IEEE/CPMT, pp.313 - 319, Oct. 1999.
- [20] Deborah S. Patterson, Peter Elenius and James A. Leal, "wafer bumping technologies- a comparative analysis of solder deposition processes and assembly considerations", *Proceedings on flipchip patterns*, Aug. 1999.
- [21] Fei Shen and Han Wang, "Real Corner Detection Based on Finding Edge Points Locally", JCIS 2002,pp. 773-776, 2002.
- [22] Harrison M. Wadsworth, "Statistical methods for engineers and scientists", McGraw-Hill Publishing Compan, pp16.2-16.22,1998.
- [23] S.Sinzinger and J.Jahns, "Implementation of an integrated confocal sensor using planar optics", *IEEE Transactions on computer vision*, 11(1), pp. 215-221, 2000.
- [24] Shabana, Ahmed A., "Computational dynamics", New York: Wiley, c1994, pp. 68-81, 1994.
- [25] Robert L. Clark, William R. Saunders, Gary P. Gibbs, "Adaptive structures: dynamics and control", New York: Wiley, c199,pp. 447-459, 1998.
- [26] Stanoje Bingulac, Hugh F. VanLandingham, "Algorithms for computeraided design of multivariable control systems", New York: M. Dekker, c199, pp.271-275, 1993.
- [27] Etienne Vincent and Robert Laganeire, "Detecting planar Homographies in an image pair", Proceeding on 2nd International Symposium on Image and Signal Processing and Analysi, pp. 182-187, Pula, Croatia, June 2001.
- [28] Alison Noble, "Descriptions of Image Surfaces", PhD thesis, Department of Engineering Science, Oxford University, p45,1989.
- [29] Schick, A. and Kedziora, M., "Inspection and Process evaluation for Flip Chip Bumping and CSP by scanning 3D confocal microscopy", Proceedings

- on 8th International Symposium on Advanced Packaging Materials, pp.116 119, March 2002.
- [30] Li, J., McIntyre, M., Lee, K. and Worster, B., "Production use of an integrated automatic defect classification (ADC) system operating in a laser confocal/white light imaging defect review station", Proceedings on Advanced Semiconductor Manufacturing Conference and Workshop, ASMC 96IEEE/SEMIpp.107 111, Nov. 1996.
- [31] 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.
- [32] Kaucic, R.; Hartley, R.; Dano, N., "Plane-based projective reconstruction", Proceedings on the Eighth IEEE International Conference on Computer Vision, vol.1, pp.420 - 427, July 2001.
- [33] Chesi, G., Garulli, A., Vicino, A. and Cipolla, R., "Estimating the fundamental matrix via constrained least-squares: a convex approach", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.24, pp.397-401, March 2002.
- [34] 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.
- [35] Moisan, L. and Stival, B., "Automatic detection of rigid point matches", Proceedings on Motion and Video Computing, pp. 235 - 240, Dec. 2002.
- [36] 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.
- [37] Sakuma, H., "Interferometry for rough surface", International Conference on Lasers and Electro-Optics Europe, pp. 230-235, Sept. 2000.

- [38] Nasrabadi, N.M., Liu, Y. and Chiang, J.L., "Stereo vision correspondence using a multichannel graph matching technique", *IEEE Proceedings on International Conference on Robotics and Automation*, Vol.3, pp.1804 -1809, April 1988.
- [39] Kuan, D., "Constraint and consistency in stereo matching", IEEE International Conference on Acoustics, Speech, and Signal Processing, Vol.11, pp.2079 2082, Apr. 1986.
- [40] Maur, F, "X-ray inspection for electronic packaging latest developments", International Proceedings Conference on Electronic Packaging Technology, pp. 235 - 239, Oct. 2003.
- [41] Ronny Theriault, "THE FMI TECHNOLOGY", SolVision, Quebec, Canada, Sep. 2002.
- [42] Mei Yu, Gang-Yi Jiang, Sai-Long He, Bo-Kang Yu and Ran-Di Fu "New approach to vision-based BGA package inspection", Submitted to Proceedings on International Conference on Machine Learning and Cybernetics, 2002., Vol.2, pp.1107 - 1110, Nov. 2002.
- [43] Sumimoto, T., Maruyama, T., Azuma, Y., Goto, S., "Development of image analysis for detection of defects of BGA by using X-ray images" IEEE Proceedings of IMTC '03. on Instrumentation and Measurement Technology Conference, Vol.2, pp.1131 1136, May 2003.
- [44] Kuk Won Ko, Young Jun Roh and Hyung Suck Cho, "A neural network approach to the inspection of ball grid array solder joints on printed circuit boards", Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks Vol.5, pp.233 - 238, July 2000.
- [45] Sumimoto, T., Maruyamay, T., Azuma, Y. and Goto, S., "A novel approach for flip chip solder joint quality inspection: laser ultrasound and interferometric system", IEEE Transactions on Packaging and Manufacturing Technology, Part A: Packaging Technologies, Vol. 24, pp. 616 624, Dec. 2001.

- [46] Sumimoto, T., Maruyamay, T., Azuma, Y. and Goto, S., "Detection of defects at BGA solder joints by using X-ray imaging", IEEE ICIT '02. Conference on Industrial Technology Vol.1, pp. 238 - 241, Dec. 2002.
- [47] Asgari, R., "Semiconductor backend flip chip processing, inspection requirements and challenges", Electronics Manufacturing Technology Symposium, 2002. IEMT 2002. 27th Annual IEEE/SEMI International, pp. 18-22, 1999.
- [48] Pyunghyun Kim an Sehun Rhee, "Three-dimensional inspection of ball grid array using laser vision system", IEEE Transactions on Components, Packaging and Manufacturing Technology, Vol.22, pp.151 - 155, April 1999.
- [49] Ho, H.M., Oh, K.E. and Wong, E.H. "Application of moire interferometry in electronics packagin", Proceedings of the Electronic Packaging Technology Conference, 1997., Vol. 1, pp.277 - 282, Oct. 1997.
- [50] Miller, M.R., Mohammed, I. and Xiang Dai, "Analysis of flip-chip packages using high resolution moire interferometry", Proceedings on Electronic Components and Technology Conference, pp.979 986, June 1999.
- [51] Rideout, E., "BGA inspection", Proceedings on Electronics Manufacturing Technology Symposium, Vol.1, pp. 388.
- [52] Deborah S, Patterson, Peter Elenius, "wafer bumping technologies-a comparative analysis of solder deposition processes and assembly considerations", Flip Chip Technologies..
- [53] Maurudis, A.S, "An efficient vector-space approach for accurately modeling fixed-point digital signal processors", IEEE Proceedings on Digital Signal Processing, Vol.2,pp.659 - 664, Nov. 1996.
- [54] Jimenez, L.and Landgrebe, D., "High dimensional feature reduction via projection pursuit", Geoscience and Remote Sensing Symposium IGARSS '94., Vol.2,pp.1145 1147, Aug. 1994.

- [55] Jin, H.,Ooi, B.C., Shen, H.T., Yu, C. and Ao Ying Zhou, "An adaptive and efficient dimensionality reduction algorithm for high-dimensional indexing", Proceedings on Data Engineering, pp.87 - 98, March 2003.
- [56] Aggarwal, C.C. and Yu, P.S., "Redefining clustering for high-dimensional applications", *IEEE Transactions on Data Engineering*, Vol. 14, pp.210 -225, April 2002.
- [57] Amir, A., Kashi, R., Netanyahu, N.S., Keim, D. and Wawryniuk, M, "Analyzing high-dimensional data by subspace validity", *IEEE Inter*national Conference on Data Mining, pp.473 - 476, Nov. 2003.
- [58] Gotze, J., Pauli, R., Schwiegelshohn, U., "Hyperbolic rotation of subspaces and VLSI algorithms for signal processing and numerical linear algebra", IEEE International Symposium on Circuits and Systems, Vol.2, pp.1604 -1607, May 1990.
- [59] Zelnik-Manor, L. and Irani, M., "Multi-view subspace constraints on homographies", Proceedings of the Seventh IEEE International Conference on Computer Vision, Vol.2, pp.710-715, Sept. 1999.
- [60] Fitzgibbon, A.W, "Simultaneous linear estimation of multiple view geometry and lens distortion", Proceedings of the 2001 IEEE Computer Society Conference on CVPR, Vol.1, pp.125-132, Dec. 2001.
- [61] Jiunn Hsuh Lim and Mani Maran Ratnam, "Optical signal processing for shape measurement", Sixth International, Symposium on Signal Processing and its Applications, Vol.2, pp.454 - 457, Aug. 2001.
- [62] Frigui, H. and Krishnapuram, R, "A robust competitive clustering algorithm with applications in computer vision", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.21, pp.450 465, May 1999.
- [63] Urfaloglu, O., "Robust estimation of camera rotation, translation and focal length at high outlier rates", Proceedings First Canadian Conference on Computer and Robot Vision, pp. 464 - 471, May 2004.

- [64] Haifeng Chen and Meer, P., "Robust regression with projection based M-estimators", Proceedings. Ninth IEEE International Conference on Computer Vision, Vol. 2, pp. 878 885, Oct. 2003.
- [65] Calafiore, G.C., "Outliers robustness in multivariate orthogonal regression", IEEE Transactions on Systems, Man and Cybernetics, Vol.30, pp.674 - 679, Nov. 2000.
- [66] Robert L. Launer and Graham N. Wilkinson, "Robustness in statistics", Proceedings of a workshop Sponsored by the Mathematics Division, pp. 61-74, 1978.
- [67] Rand R. Wilcox and Huber, "Introduction to robust estimation and hypothesis testing", Academic Press, pp.12-16, 1997.
- [68] Rami S. Mangoubi, Michael J. Grimble, "Robust estimation and failure detection", Springer-Verlag, pp. 124-128, 1998.
- [69] D.R. Myatt, P.H.S. Torr, S.J. Nasuto, J.M. Bishop, and R. Craddock. Napsac, "High noise, high dimensional robust estimation - it's in the bag", British Machine Vision Conference, Cardiff, UK, pp.458-467, 2002.
- [70] O. Chum and J. Matas, "Randomized ransac with t(d,d) test", British Machine Vision Conference, Cardiff, UK, pp. 448-457, 2002.
- [71] Okabe, T. and Sato, Y., "Object recognition based on photometric alignment using RANSAC", Proceedings IEEE Computer Society Conference on Computer Vision and Pattern Recognition Vol.1, Pages:I-221-I-228, June 2003.
- [72] A. J. Lacey, N. Pinitkarn and N. A. Thacker, "An Evaluation of the Performance of RANSAC Algorithms for Stereo Camera Calibration", Proceeding on BMVC 2000, section VI-Multiview Technique, Sep. 2000.
- [73] 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.

- [74] Michael Werman and Yoram Gdalyahu, "Ridge's Corner Detection and Correspondence", Proceedings on Conference on Computer Vision and Pattern Recognition, pp. 976, Jun 1997.
- [75] C.G. Harris and M.J. Stephens, "A combined corner and edge detector", Proceedings Fourth Alvey Vision Conference, Mancheste, pp 147-151, 1988.

