Fast measurement of micro-bumps on chip package using the coaxially optical system: low-coherence interferometry

Joonho You[1], **Seung-Woo Kim** [2,#]

1 Intekplus Co. Science Town, Daejeon, 305-701, South Korea
2 Department of Mechanical Engineering, KAIST, 335 Gwahak-ro(373-1 Guseong-dong), Yuseong-gu, Daejeon, South Korea, 234-567
Corresponding Author / E-mail: swk@kaist.ac.kr, TEL: +82-42-350-3217, FAX: +82-42-350-5217

KEYWORDS : Micro-bumps, Low-coherence interferometry, Fast measurement

Low-coherence interferometry is suitable for 3D profiling demands in diverse industries such as semi-conductor, flat panel display and optics. In this paper, we focus on 3D measurement for micro-bumps in chip package industry which contain various optical properties, for example, long step heights, rough surface, round shape and large gap of reflectance on a surface in a field of view. These combined optical properties make it difficult to measure the surface of target with conventional microscopic interferometry. To get over barriers for the measurement of micro-bumps, Twymann-Green configuration has been adapted into the proposed system instead of using the interference microscopic lens and Detailed analysis for the properties has been done in terms of how we enhance the visibility of an interferogram. According to the analysis, we successfully designed specialized measurement system which is dedicated for micro-bump height inspection. It has dual reference mirrors for long step height, various reflectance for optical coating on the mirror surface, electrical synchronization system for fast measurement and imaging system for round-shape measurement. Finally, we have verified how it works successfully and implemented wafer inspection machine on which the proposed optical probe is attached.

Manuscript received: January XX, 2011 / Accepted: January XX, 2011

1. Introduction

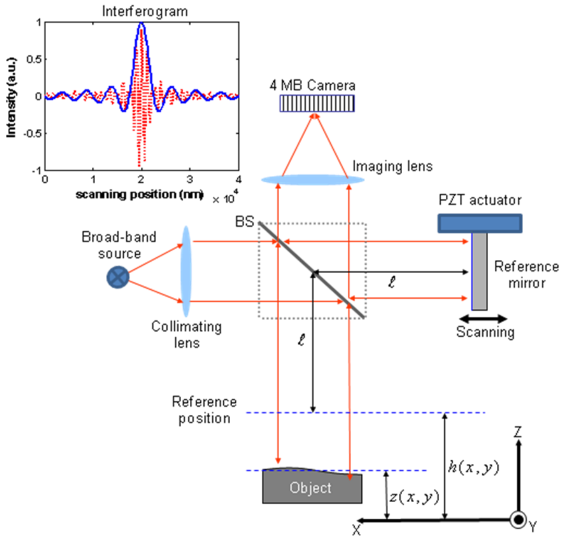
Low-coherence or white-light scanning interferometry is a sophisticated technology for industries, such as MEMS, semiconductors, and flat panel displays, which require high precision measurement technology. Moreover, large-area and real-time surface measurement has been strongly demanded for an in-line measurement in these fields. In implementing the three-dimensional, large-area, and real-time measurement, one of the most significant technologies is the real-time processing of the large amount of interference signals, which come from a large-size imaging sensor. Several researches for high-speed and larger-area measurement have been studied so far. For instance, Dresel et al implemented the Twyman-Green type white-light scanning interferometry instead of the microscope type for measuring large-area metal surfaces[1]. Hirabayashi et al proposed a specialized algorithm, squared-envelope function estimation by sampling theory (SEST), which can extend scanning step up to 1 μm while keeping the precision up to 10nm[2]. Fleischer et al improved the computational performance by several times in white-light scanning interferometry by using the multi-media extended (MMX) technology of an Intel central processing unit (CPU) [3].

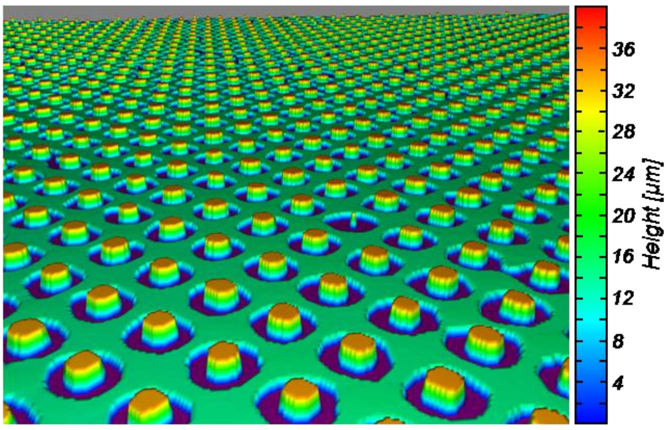
In this paper, we describe a real-time low-coherence scanning interferometry for large-area measurement of PCB micro-bumps by using the parallel processing technology with the GPU. Four megabyte resolution and a high-speed CMOS camera are used, and during capturing low-coherence fringe images, Optimized parallel processing method finishes the calculation of a dedicated centroid algorithm with Four megabyte images within the interval of grabbing the fringe images. In the end, total measurement speed is equal to the camera speed without any post processing, regardless of how many frames are applied.

2. Optical system and measurement algorithm

2.1 Optical system

Both an optical setup and an algorithm are dedicated to the real-time profiler for high-speed and large-area measurement. Twyman-Green configuration is adopted as the basic schematic of the low-coherence scanning interferometry for large-area imaging, and a weighted recursive centroid algorithm is used into parallel processing with the GPU.





(a) Optical system (b) Measurement result of PCB bumps

Fig. 1 Optical system and t measurement result of PCB bumps

We implemented the parallel processing for the centroid algorithm through the Graphic Processing Unit(GPU), which is designed by NVIDIA. The GPU has up to 1,000 stream processors which are dedicated to 32-bit floating-point operation, such as addition and multiplication. We also used the Nvidia's Compute Unified Device Architecture (CUDA) as a software platform for enormously parallel high-performance computing on its GPU

2.2 Measurement algorithm

The low-coherence interferogram is regarded as the incoherent superposition of individual interferogram of all monochromatic waves that form a wide-spectrum light source. Accordingly, the model of the white-light interferogram sampled from the scanning of the reference mirror is described in the form of the integral of the wide-spectrum. From the viewpoint of the fast processing of the large interferogram from a four megabyte CMOS camera, there are two critical requirements. Firstly, the structure of the algorithm should be recursive in order to finish the calculation of the algorithm during the scanning and reduce a lot of memory-consumption by many grabbed fringe images. Secondly, the algorithm should be simple for fast processing and robust to noises which come from external vibration and systematic error. To do so, we adopted the weighted recursive centroid method detecting the maximum peak of the envelope function, which is perfectly bell-shaped and symmetric to the maximum peak. After getting fringe images, we convert raw signals into the absolute of first-order derivative of the grabbed interferograms, a proper threshold is applied to eliminate noises and then the weighted centroid is calculated recursively. The maximum peak of the envelope function is exactly equal to that of the absolute of first-order derivative of the interferogram.

$$h(x,y)=\frac{\int\left[w(z)\left|I'(z)\right|\right]zdz}{\int\left[w(z)\left|I'(z)\right|\right]dz} \tag{1}$$

4. Conclusions

In this paper, we implemented and verified a real-time surface profiler using a low-coherence scanning interferometer accelerated by the parallel processing of the GPU. For large-area and high speed measurement, Twyman-Green type was adopted as an optical system. Even though we used a large-area and fast CMOS camera, we successfully eliminated the post-processing by using the GPU. Through comparison, we proved that the GPU is 160 times faster than the CPU. From this research, in terms of large-area and real-time measurement, we showed that the speed of a camera would decide the measurement speed of low-coherence scanning interferometry.

REFERENCES

1. Dresel T., Häusler G., and Venzke H.(1992) 'Three-dimensional sensing of rough surfaces by coherence radar', Appl. Opt. 31, 7, 919-925

2. Hirabayashi A., Ogawa H., and Kitagawa K. (2002) 'Fast surface profiler by white-light interferometry by use of a new algorithm based on sampling theory', Appl. Opt. 41, 23, 4876-48830.

3. Fleischer M., Windecker R., and Tiziani H. J. (2002) 'Fast algorithms for data reduction in modern optical three-dimensional profile measurement systems with MMX technology', Appl. Opt. 39, 8, 1290-12977.