nDSE Based Overlay Alignment: Enabling Technology for Nano Metrology and Fabrication

Jun Gao[1], Carl Picciotto, Wei Wu, Inkyu Park, William M. Tong
HP Labs, 1501 Page Mill Road, Palo Alto, California 94304 USA

ABSTRACT

Displacement sensing and estimation (DSE) is important preprocessing task for many image-based processing systems that extract information from multiple images. In last two years, we gained significant insight of the nature of DSE and developed theory and algorithm framework named *n*anoscale *d*isplacement *s*ensing and *e*stimation (nDSE). We also build procedures to apply nDSE to overlay alignment down to the nanoscale. We will introduce two basic theories: Phase Delay Detection (PDD) and Derivatives-based Maximum Likelihood Estimation (DML) and associated DSE algorithms, noticeably Near-Neighbor-Navigation (N-Cubed) algorithm. We presented our best nDSE experimental result of 1 nm (1σ) while tracking 5 nm stepping. To develop nDSE-based nanoscale alignment, we introduced our definition of displacement, alignment and pseudo-displacement. We presented both theoretical and practical procedures to use nDSE to achieve nano-alignment down to the 10s of nano-meters and beyond. Then we compared nDSE-based nano-alignment to other industry standard alignment method and attempt to show the substantial advantages of nDSE based alignment in terms of cost and simplicity of the system design.

Keywords: nDSE, IDMA, DDMA, Nano Imprint, Alignment, decentralized fabrication

1. Introduction

Displacement sensing and estimation (DSE) is important preprocessing task for many image-based processing systems that extract information from multiple images. However, DSE is not generally considered applicable to overlay alignment, especially down to the nanoscale. This is essentially due to a general mis-perception that optical displacement sensing is fundamentally limited by the wavelength of light, optical point-spread function and sensor pixel size (Rayleigh and Shannon/Nyquist theorems). In last two years, we gained significant insight of the nature of DSE and developed theory and algorithm framework named *n*anoscale *d*isplacement *s*ensing and *e*stimation (nDSE). We also build procedures to apply nDSE to overlay alignment down to the nanoscale.

We developed algorithms, based on two fundamental theories: *p*hase *d*elay *d*etection (PDD) and *d*erivatives-based *m*aximum *l*ikelihood estimation (DML). PDD models displacement as phase delays of the target system, thus providing a foundation for frequency-domain DSE; while DML defines the principle of displacement estimation from the perspective of minimizing error propagation in the maximum likelihood sense between observation system and the target system.

The algorithm we used in our research is a correlation based algorithm: *N*earest *N*eighbor *N*avigation algorithm (N-Cubed). We derived an analytical solution for displacement estimation that models an arbitrary correlation surface as the general 2nd order 2-Dimensional Taylor expansion. We formulated the general linear least-squares problem by comparing data in correlation grid with the expected values computed using he second-order model. We then compute the optimum values of the parameters by minimizing analytically the corresponding sum of the squared errors, obtaining a solution set consisting of simple linear combinations of the correlation data values. N-Cubed algorithm is an extremely robust and efficient algorithm. It provides a clean and simple analytical solution without iteration; and is quite tolerance in the presence of white noise. In ideal conditions, it can provide displacement accuracy down to 1/200th of a pixel dimension or better. With pixel size of 100 nm, (50x lens and a 5 micron pixel array) N-Cubed algorithm achieved, experimentally, step size of 5 nm displacements tracking with one sigma precision of 1 nm.

We discuss the application of nDSE as an overlay metrology tool. The key is to bridge DSE and alignment measurement. We presented definitions for displacement and alignment and from there; we presented several methods to bridge these two concepts by clever selection of references for DSE. DDMA (*d*irect *d*isplacement-*m*easurement-

[1] Jun Gao: email: jun.gao@hp.com; Phone: 1+650-857-5050; Fax: 1+650-857-4320; HP Labs, Palo Alto, CA USA

Metrology, Inspection, and Process Control for Microlithography XX, edited by Chas N. Archie, Proc. of SPIE Vol. 6152, 615223, (2006) · 0277-786X/06/$15 doi: 10.1117/12.656610

based *a*lignment) is the first method we used to generate reliable alignment results using nDSE. The best result we achieved was 21nm alignment accuracy using N-Cubed algorithm. IDMA (*i*ndirect *d*isplacement-*m*easurement-based *a*lignment) is the most accurate in theory, and hardest to prove experimentally. It is promising because it truly utilizes DSE to achieve precision co-placement over different layers and processes. IDMA and DDMA, as well as a group of other nDSE based algorithms/procedures are vibration tolerance, which makes them ideal as an overlay alignment tool for nanoimprint lithography.

Current methods of overlay metrology and many methods of displacement metrology require precise alignment targets, such as symmetric geometric figures or extremely high-Q diffraction gratings. Such patterns are expensive to produce and/or difficult to fabricate consistently and they occupy valuable real estate on the wafer. nDSE based alignment framework provides precision alignment by tracking totally arbitrary patterns. We extend this advantage into a method for alignment sensing, which retains nDSE as the key underlying measurement. Hence, as with displacement sensing, the alignment targets need not be held to any absolute standard, pattern asymmetries caused by process variations are not an issue, and precision gratings are not required. These advantages of (D)IDMA are summarized below. We believe these advantages will make nDSE based alignment one of the enabling techniques for nano metrology and fabrication.

2. Displacement sensing and estimation to the nanoscale

2.1. nDSE theory: Phase Delay Detection

In frequency domain, displacement committed by a rigid object is equivalent to a constant group delay of all observed frequencies. The general phase correlation method (PCM) is a well known technique in estimating displacements, or estimating general transformations such as affine transformations [1] [2] [3]. PCM has also been extended into sub-pixel accuracy with varying levels of success [3] [4] [5] [6] [7] [9]. It has also been used to estimate rotations [8]. Compared to spatial-correlation-based algorithms, PCM can be made particularly robust in the presence of band-limited noises and distortions. Since the phase difference for every frequency contributes to the resulting displacement estimate, the location of the peak, which represent the displacement, will not change if there is noise which is limited to a narrow bandwidth [3]. Consequently PCM is particularly effective when processing images with illumination differences, since those differences are usually slow varying and therefore concentrated at low-spatial frequencies [1]. On the other hand, PCM offers little solution to deal with white noise which, since by definition it's spread across all frequencies, cannot be removed. The location of the phase correlation peak will be inaccurate since the phase difference at each frequency is corrupted [3].

Phase Delay Detection (PDD) theory and algorithms are closely related to the Phase Correlation Method (PCM). They both treat displacements from a reference frame to a comparison frame as phase differences in the frequency domain. Although other correlation-based algorithms might use Fourier Transforms (FT) in their operation, they differ from PCM-based algorithms because PCM-based algorithms search for the optimal match according to information in the frequency domain, while other algorithms might use the Fourier Transform as a tool to perform a spatial operation, for instance, spatial correlation.

The major difference between PDD and the majority of PCM-based algorithms lies mostly in the last conceptual step of determining the displacement from the phase correlation information. While most other PCM-based algorithms rely on the inverse Fourier Transform to determine the displacement, PDD extracts the displacement directly from the phase response in the frequency domain. The deeper theoretical difference between PDD and most other PCM methods lies in the method of extracting the degree of coherence between reference and comparison images information. C. D. Kuglin, etc (1975) [1] suggested that "*the phase correlation algorithm is based upon the fact that the information pertaining to the displacement of two images resides in the phase of the cross power spectrum*", and this principle is the foundation of most PCM-based algorithms [3]. PDD treats displacement as a phase delay, just as other PCM algorithms. However, it also treats a displacement as the application of a "shift operator", or a linear shift-invariant system with constant group delay. Therefore, the phase correlation information resides in the transfer function of this LSI system. It does not perform "whitening", or cross-normalization between reference and comparison frames. PDD actually preserves the phase information, along with its noise sources and distortions. It then "masks-out" these noise sources and distortions according to magnitude information in frequency domain or other prior knowledge about the noise/distortion sources.

The final displacement or registration estimate is achieved without leaving the frequency domain by least-squares fitting the phase-difference plain after the regions of noise; distortion, aliasing, illumination non-uniformity, etc… are masked out.

The concept of extracting the displacement directly from phase differences was mostly discounted [4] for been too noisy. Stone and Orchard (1999) [6] proposed a method similar to PDD to deal with first-order aliasing issues with some success. What we propose here is a systematic method to mask out noise sources and distortions so that PDD can potentially perform better than the general PCM algorithm.

PCM-based algorithms all take advantage of the shift property of Fourier Transform. The shift (transformation, delay, etc…) property in the 2-D case is defined as follows,

$$f(x-x_0, y-y_0) \Leftrightarrow e^{[i2\pi(ux_0+vy_0)]}\mathsf{F}\,(u,v) \tag{1}$$

Where x_0 and y_0 are shifts (displacements) in x and y directions, respectively, and $F(u,v)$ denotes the Fourier transform of a 2-D function (image) $f(x,y)$.

Displacement is represented by operation of a Linear Shift-Invariant System, as depicted by Figure 1. The shift operator imposes a constant group shift (or group delay, as it is termed in communications theory), thereby exhibiting linear phase, as evidenced by the equality of equation (2) which results directly from equation (1). Equation (2) shows directly that the operation of a shift operator results in a linear phase shift, with phase shifts varying linearly with frequency.

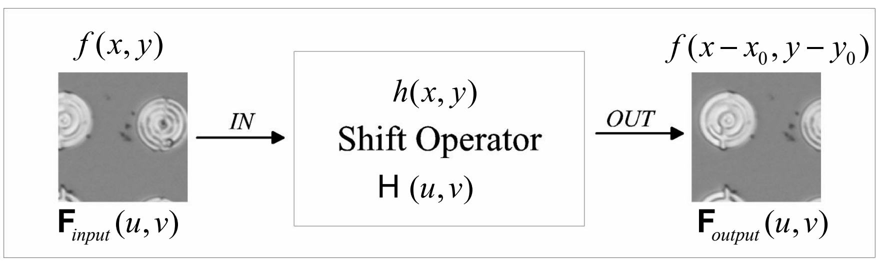

Figure 1: Phase Delay Detection Operation as a Linear Shift-Invariant Operator

$$\mathsf{H}\,(u,v) = \frac{\mathsf{F}_{output}(u,v)}{\mathsf{F}_{input}(u,v)} = e^{i2\pi(ux_0+vy_0)} \tag{2}$$

The mathematical difference between general PCM-based algorithms and PDD is that the general PCM-based algorithm uses the cross-power spectrum of the two images to derive the phase difference:

$$\frac{\mathsf{F}_{input}(u,v)\mathsf{F}^*_{output}(u,v)}{\left|\mathsf{F}_{input}(u,v)\mathsf{F}^*_{output}(u,v)\right|} = e^{i2\pi(ux_0+vy_0)} \tag{3}$$

Where F^* is the complex conjugate of F.

The phase of the cross-power spectrum is equivalent to the phase difference between the images, as used by PDD (equation (2)). PCM-based algorithms rely on the inverse Fourier transform to locate the peak of the correlation. The inverse Fourier transform generates a delta function with an offset corresponding to the displacement. The method is generally very efficient when only pixel-level displacement accuracy is required. There have been numerous attempts to extend general PCM algorithms into sub-pixel accuracy [1] [2] [3] [4] [5] [6]. These algorithms more-or-less spread the peak from a single pixel into multiple, surrounding pixels by up/down sampling, and then estimate the more finely resolved location of the peak by interpolation.

We propose to locate the displacement without leaving the frequency domain. For this purpose, consider a 3-D space with its X-Y reference frame given by the two frequency axes u and v. The phase difference of the reference frame and the displaced comparison frame is the third axis. In this space, $ux_0 + vy_0 = 0$ defines a plane through the origin whose slope along the two frequency axes specify the shifts along the two spatial axes. Because of noise and distortion in and between the reference and comparison frames, the plain defined above would be quite noisy and inaccurate in some frequency areas. However, as long as we retain at least three good, noise-free, undistorted points, it should be possible to fit and reconstruct a plane. Of course, it is generally beneficial to retain as many points as possible. With additional frequency-specific information, such as the magnitude response of each individual frame or the magnitude response of the LSIS $|H\ (w)| = |\mathbf{F}_{output}(u,v) / \mathbf{F}_{input}(u,v)|$, most of the noisy and/or aliased frequency bands can be eliminated, or masked-out. For example, by examining the magnitude response of the individual reference/comparison frames, frequency bands with magnitudes smaller than a pre-set threshold can be masked-out for being too noisy, possessing SNR too small to extract meaningful phase relationships. Rules can also be developed to mask-out frequency bands which would produce numerical problems in phase computation. By noting deviations from unity in the magnitude response of the ratio, Stone Harold etc. showed that first-order aliasing can be effectively reduced by masking [6].

Other prior knowledge both about the target frames and the observation system can be used to generate masks. Frequency bands of patterned sensor noise, optical distortion, illumination issues, etc, can all be determined beforehand and masked-out so that those frequency bands will not corrupt the estimates of the phase slopes. Finally, strong frequency bands known to have high SNR can be given higher weighting. Occasionally it may be desirable to design target patterns with enhanced specific frequency bands for this purpose.

One issue to consider is the difficulty of sensing large displacements, for example displacements greater than the wavelength of the higher frequency components of the target image. In such cases, care should be taken to use only the lower frequency bands, since the higher frequency bands may be subjected to phase "wrap-around" past 2π radians. In some applications, it might be appropriate to use the lower frequencies to achieve coarse displacement estimation, and then, once the reference and comparison frames are shifted to within close proximity, to use the entire frequency range to achieve fine displacement estimation.

2.2. The limits of DSE: observation system and target system error propagation model (DML)

In this section, we provide error analysis and an error propagation model for nDSE algorithms based on Derivatives-based Maximum Likelihood Method (DML). We also pursue deeper insight by providing an information system perspective for displacement sensing and estimation. We labor towards a model linking the performance of the observation system and the target system's information content with the maximum accuracy theoretically achievable by any displacement sensing and estimation algorithm. We also derive a close relation between the real-spaced derivative methods and the Fourier-spaced methods. More research in this area is needed in the future to provide a more complete theory, relating how the observation and target systems govern the maximum achievable displacement sensing and estimation performance. Readers interested into the mathematical modeling and detailed mathematical deduction should refer to reference [12] "Displacement sensing and estimation theory and application", by J. Gao and others, published in Applied Physics A in March 2005.

The nDSE error analysis is an important part of our theory research. The goal of the error analysis is to find the theoretical limitation of DSE, and the relationship between the capabilities of the observation system, which records the displacement, and the characteristics of the object, which expresses the displacement. The theoretical model we use is based on the Derivative-based Maximum Likelihood method [12]. In the general case, it is mathematically complicated and beyond the scope of this paper. However, if we can make an extreme simplification and assume that the object image is bandwidth limited and only has vertical and horizontal features, the theory maintains that the error of displacement estimation is inversely proportional to the size of the surface-feature derivatives and the square root of the number of pixels with non-zero derivatives in the direction of the displacement. The total error in the estimate is reduced by a factor of the square root of the number of pixels, a result consistent with the reduction in noise of an average of a number of independent measures averaged to find the joint estimate of the mean. From these results, the minimum error is found for patterns which maximize the derivative over the most number of pixels. These error analyses points to DSE's major advantage over other methods, which is that it only requires that the object have some features, ideally features with derivatives maximized along expected displacement direction. It does not require specially designed targets with regular patterns, from either a spatial-domain or a frequency-domain perspective. This advantage, and the fact that far-field optic microscope observation systems can be used to achieve nanoscale displacement sensing, make nDSE a powerful enabler for the application framework IDMA/DDMA, which will be discussed in the following section.

2.3. Correlation-based nDSE algorithms

nDSE algorithmic implementations generally fall into two groups: (1) algorithms based on statistical analyses of images, such as correlations of the reference (non-displaced) frame with respect to the comparison (displaced) frame; (2) algorithms based on transforms, such as the Fourier transforms of the individual frames. Algorithms of the latter type, such as PCM (Phase Correlation Method) and PDD (Phase delay detection), take advantage of the translation property of Fourier transforms, as expressed in equation (1).

Our research of nDSE and DSE-based alignment by and large involved implementations of algorithms based on statistical analysis. One particular implementation is named *N-Cubed* algorithm, which stands for *N*earest *N*eighbor *N*avigation algorithm. The name came from implementations of this algorithm in recent HP micro-scale navigation applications; such as HP's successful optical media advance sensor and handheld scanner. This algorithm, as shown in Figure 2 below, involves multiple 2-dimensional correlations between the two images. The correlation function we most often use in N-cubed based algorithms is defined as

$$C_{i,j}^{k} = \sum_{m=1}^{M}\sum_{n=1}^{N}\left|r_{m,n} - c_{m-i,n-j}\right|^{k}, \tag{3}$$

A non-zero-mean difference-detection function, where $r_{m,n}(c_{m,n})$ is the digitized value of the reference (comparison) image at pixel $\{m,n\}$ and $i, j \in R$ represent the shift of the comparison frame relative to the reference frame, measured in units of pixels. Typically, $k \in \mathrm{N}$, although in principle we can have any $k \in R \geq 1$[12]. In all our experimental implementations of the N-Cubed algorithm, we always chose $k = 2$, resulting in a more intuitive measure which computes the sum of the squared pixel differences between the reference frame and the comparison frame (refer to figure below). After the raw correlation surface of 3x3 or 5x5 region is acquired, a general 2nd order Taylor expansion is fitted to these correlation figures and chi-square error between the raw correlation surface and the fitting function is minimized to yield the coefficients for the Taylor expansion, which is then used to compute the sub-pixel accuracy displacement between reference and comparison frames. This implementation provides a closed analytical solution of displacement estimation, which is both efficient and elegant. The computation complexity is bounded and no iteration is required. In most cases, the speed of update of displacement information was limited by the image acquisition system, rather than N-Cubed algorithm.

2.4. nDSE experimental results

We experimentally investigated the practical limits to the precision of our nDSE algorithms. We first reported our findings here [11]. We set-up an optical microscope with a 4.8-micron CCD camera and a 50x objective lens, resulting in a pixel resolution of approximately 96nm. We used broadband white light, so the optical resolving power was clearly less than the pixel resolution. Appropriately, from a sampling theorem (Shannon/Nyquist) standpoint, we were over-sampling somewhat. Under the microscope we placed a piezo actuator stage manufactured by nPoint. We used the stage to translate fragments of arbitrarily-patterned silicon (Fig. 3) in 5nm steps, dwelling for approximately ten seconds at each step, while we captured a continual stream of images. We then processed the CCD images using nDSE and plotted the results (Fig. 4). The images were taken at approximately video rates (approx. 20 frames-per-second), and no averaging was performed. Each point in the graph of Fig. 4 corresponds to a single image. Cleary, we can easily discern 5nm steps using nothing more than an optical microscope and nDSE. We calculated a measurement standard deviation (one σ) error of approximately 1nm, due in part to positional uncertainties caused by vibrations in the apparatus.

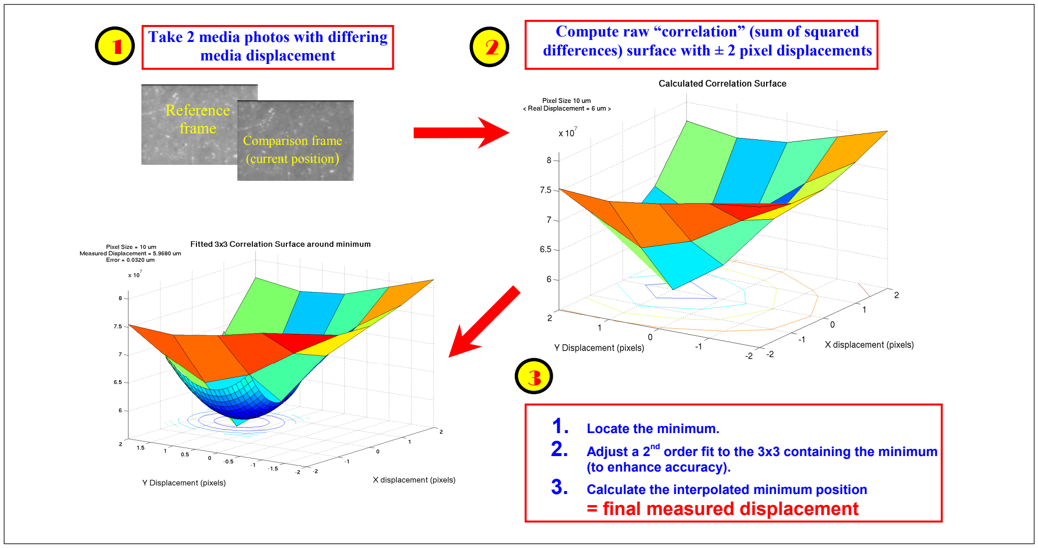

Figure 2: Spatial domain correlation based *N*earest *N*eighbor *N*avigation (N-Cubed) algorithm

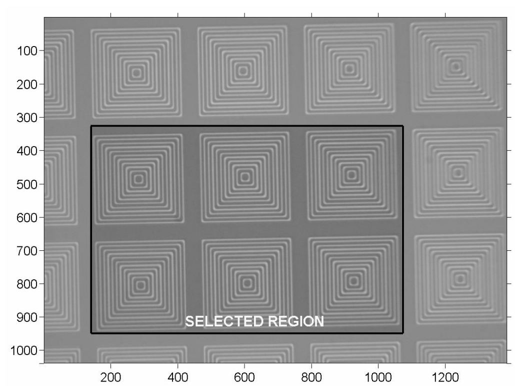

Fig. 3: The patterned silicon wafer fragment we tracked on the piezo stage

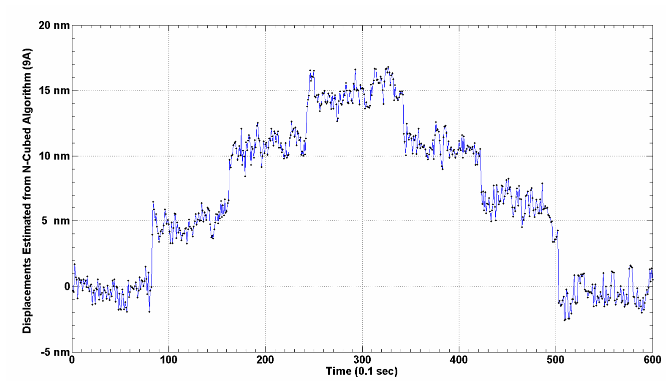

Fig. 4: Demonstrated accuracy of nDSE. Each step is a real step of 5nm, as reported by our nPoint piezo stage

Note that although we employed a far-field optical observation system in our nDSE research, the nature of observation system is not limited to optical system. We experimented with SEM images and other similar imaging systems, which can provide similar degree of accuracy relative to the measurement element of that particular system. Our DSE theory and algorithms can work with any observation system which can provides data matrix representative of the location of the target system, optical or non-optical; 2-dimensional or higher dimensional. This makes our DSE theory and algorithm power, versatile and flexible in many application areas.

3. From nanoscale Displacement Sensing & Estimation to Nanoscale Alignment

3.1. A set of definitions

In order to fully explain DSE-based alignment procedures, it is necessary to provide a clear definition of two key concepts in this discussion: displacement and alignment. From the definitions, we can explore the link to bridge the differences between them from their common properties.

1. **Displacement:** A vector or the magnitude of a vector from the initial position to a subsequent position assumed by an object
2. **Alignment:** Multiple objects have pre-determined distances from common reference
3. **Pseudo-displacement:** "displacement" between different objects made to be as similar as possible

The definition of displacement is straight forward. However, it does reveal some of the characteristics that separate displacement from alignment. Firstly, displacement is about ONE object; Secondly, there is a hidden time reference, which means that the initial position assumed by the object happen BEFORE the subsequent position assumed by the same object; and thirdly, displacement is a relative measurement. It is about the location of the SAME object relative to previous location. Alignment definition is slightly expanded from the industry standard center-line based overlay concept (Alexander Strarikov [10]). Our broader definition of alignment has essentially three characteristics: (1) alignment involves two or more objects; (2) there is no hidden time sequence, unlike the displacement. The objects distance can be determined without a defined sequence; and (3) alignment is about "pre-determined distances" to a common REFERENCE between objects. The distance does not necessary need to be zero, or close to zero, as desired by conventional alignment techniques. We claim that as long as the distance among objects is determined, they are aligned. This broader alignment definition is introduced to enable our DSE algorithms to achieve alignment. In our theories and algorithms, we do not require any regular or high accuracy alignment mark. The whole object under the observation system, in most cases a far field optical observation system, is used as alignment marks to achieve micro- and nanoscale DSE and alignment. In most cases, definitions of center line, the center of the alignment marks or the line width of the alignment feature have not meaning in our paradigm.

Displacement, by definition, is not alignment. However, good DSE procedures, like our N-Cubed algorithm or PDD based algorithm, can easily be used to achieve *consistent co-placement*. Meaning that using DSE based procedures, one can place the same object to the same location, relative to previous location of the object, with high precision. If a common reference can be identified and the distances between objects over multiple instances can be determined, a DSE based alignment can be achieved: *Alignment* = *Consistent co-placement* (DSE) + *Reference*.

In practice, during lithography and nano-imprint process, there are layers of patterns to be overlayed onto the substrate by using masks with different patterns. Difficulty arises when different masks needed to be placed consistently with regard to each other and substrate. A pure DSE based method could be inadequate. Pseudo-displacement concept is introduced to allow two or more different object to determine their relative positions by determine their "displacement" as if they are all same object moving from one location to another. The accuracy of this approach deteriorates from the pure DSE and is determined by how closely these objects used for DSE resemble each other. Currently, the criteria are quite empirical and experimental. We can, however, generally divide pseudo-displacement into three groups, which provide some guideline as to how accurate the pseudo-DSE would be.

True displacement sensing and estimation with same object	One object moving from initial location to subsequent location
	DSE accuracy: extremely high with right DSE algorithm; can achieve $1/200^{th}$ pixel dimension or better
Pseudo-DSE 1	Multiple objects made with same processes on the same material
	DSE accuracy: high with right DSE algorithm; can achieve ~$1/100^{th}$ pixel dimension
Pseudo-DSE 2	Multiple objects made with same processes on different material; or different process on same material
	DSE accuracy: mid to high with right DSE algorithm; can achieve $1/50^{th}$ pixel dimension
Pseudo-DSE 3	Multiple objects made with different processes on different material
	DSE accuracy: low to mid with right DSE algorithm; can achieve $1/20^{th}$ pixel dimension

Note that with different grade of PsDSE, if the processes of making the displacement marks are well controlled and the systematic errors of displacement sensing and estimation can be measured by independent metrology procedures, the precision of PsDSE can be increased, based on these metrology and calibration processes.

3.2. Linking displacement to alignment: defining a reference

Traditionally, lithographers judge mask-to-wafer alignment by overlaying complementary alignment patterns over one another until they appear to line-up in the field-of-view of an optical microscope. This apparently simple statement contains the important implication that the alignment patterns must possess centerpoints, or defining edges, which can be compared directly against one another to deduce relative positioning. Displacement sensing, on the other hand, is a relative measurement, and makes no such absolute localization determination. Rather, true displacement sensing (as opposed to pseudo-displacement sensing, discussed above in section 3.1) tracks relative spatial translations of an object over time. With such a measurement we may determine with great precision that, for example, a silicon wafer has been translated 5nm to the right, but with no further definitions we can not deduce an absolute position, nor can we perform differential position (e.g. alignment) measurements between marks.

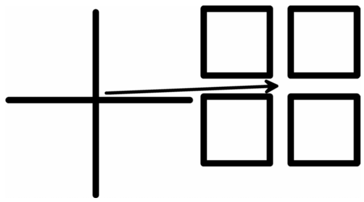

Fig 5a: Distance between alignment marks is well defined when using alignment patterns with defined centerpoints. Features are located using edge-detection and geometric extractions.

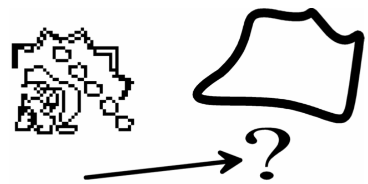

Fig 5b: Distance between alignment marks has no definition when using arbitrary features with no defined centerpoints. Individual features are tracked using relative displacement sensing.

Any absolute measurement requires an origin reference, even those using symmetric patterns with defined centerpoints. Of course, many alignment measurements are inherently differential, so the origin will subtract itself out. But, for practical purposes, we do need a defined origin as we build-up to an alignment measurement. For the case of displacement measurements, the origin can not be a specific coordinate in space, because we would be operating with no definitions for absolute position. Rather, the origin may be defined as an initial configuration of materials, such as the co-placement of a wafer and a mask. We may prefer to use the term “reference” rather than “origin” to distance ourselves from the concept of an absolute coordinate system, and to remind ourselves that we have no definition for absolute position.

With the reference image acting as our “origin,” we have transformed measurements of the translations of individual objects into measurements of differential movements between the objects. These differential movements indicate changes to their relative positions, or to their *co-placements.*

We speak of references in two contexts here. The image of Fig. 6a in its entirety acts as a reference frame. But also, we can view the image of a single object as a mechanism for tracking an origin. For example, if the wafer is considered “ground zero,” we can track changes to the wafer’s origin by following a mark on the wafer. While we still have no absolute definition of the origin’s location, we can track its relative location over time.

The purpose of this section is to introduce and emphasize the importance of the concept of a reference to transform individual relative displacement measurements into measurements of relative positioning between objects, thus enabling alignment measurements. Our alignment-from-displacement methodologies, discussed below, employ this essential concept in various manifestations.

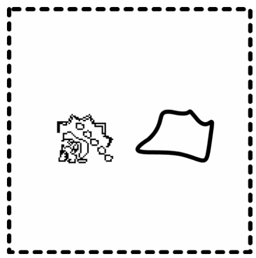

Fig 6a: A reference configuration, captured by an imaging system. We call this "origin image" a "reference frame." The two arbitrary marks are viewable simultaneously, though they physically reside on separate materials, e.g. a wafer and a mask

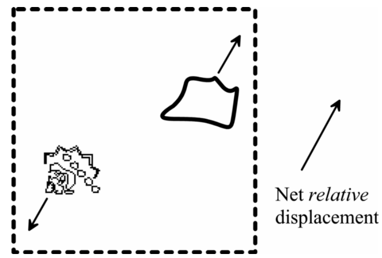

Fig 6b: A configuration at a later time. The reference frame is used as a comparison to derive a net change to the relative positions of the pieces of material.

4. Two nanoscale alignment frameworks: DDMA & IDMA

In Sections 2 and 3 we discussed the displacement-sensing-based tools at our disposal for applying to alignment sensing applications. We now discuss the frameworks by which we may practically implement these tools. We essentially lump the frameworks into two classes: DDMA and IDMA, for *Direct* and *Indirect Displacement-Measurement-based Alignment*, respectively.

In ideal implementations of DDMA and IDMA, we ensure that all alignment marks remain individually separable, meaning we do not allow the marks to overlap, so that the algorithms may crop-out and process the marks individually. If there is to be overlap, the bests cases involve marks which, when aligned, nest inside one another such that the two may still be individually extracted.

4.1. DDMA, based on pseudo-displacement estimation

DDMA describes the direct use of displacement estimation between nominally identical features to determine relative positioning between substrates or logical layers on a single substrate. Essentially, DDMA is a direct implementation of one or more of the pseudo-displacement estimation techniques described in Section 3.1, and represents the simplest approach to applying correlation processing to alignment sensing. To implement DDMA, image frames containing nearly identical (though not necessarily regular) alignment features from each of the substrates or layers are acquired and correlated against each other to estimate relative positioning. Due to variations caused by processing such as deposition, etching, or polishing, the two frames (i.e. from two separate images or from two sub-windows within the same image) typically contain alignment features which are similar but generally not quite identical. Compounding the issue, the various marks tend to be created using different processing, for example mask or mold features versus those on processed silicon. These variations lead to measurement errors. In the best cases these errors are systematic, but there is usually an unavoidable random component. It is for this reason that DDMA-only solutions have difficulty meeting nanoscale precision requirements. However, when DDMA can meet the precision requirements, DDMA provides us with a method that is relatively straight-forward and easy to understand and implement. It is the closest to the conventional alignment method in that it straightly relies on pseudo-DSE and compares the near "identical" features across material and process boundary. Essentially the accuracy is determined by the class of pseudo-DSE used and if the error is systematic and removable.

We studied DDMA experimentally first exactly because of its simplicity. The detailed technical information about the experiments; experimental results and the analysis were first published in JVST November/December issue for EIPBN conference in 2005[13].

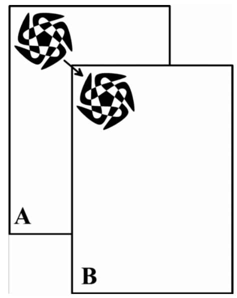

Fig 7: A pseudo-displacement measurement (pseudo-DSE) supporting a DDMA measurement between layers A and B. Each layer is patterned with nominally the same feature.

In this experiment, we achieved 81nm accuracy over 23,300 nm offset in one direction; and 21nm accuracy over 8500 nm offset in another direction. In figure 8, the X and Y axis are in unit of pixel with pixel size of about 95 nm. The alignment feature we used is a pseudo-random pattern of micro- size line width, so that the far-field microscope can clearly resolve them. They are on different material (mold and substrate) and made with different processes (refer to fig. 4 in [13]), and therefore, they are Pseudo-DSE 3 as defined in section 3.1. The experimental results shown at this level of Pseudo-DSE level, the accuracy can not be expected in nano-meter range, rather in 10s nanometer range, depends largely on if and how to control systematic error. We feel DDMA definitely has it place since we can achieve 20 to 80nm alignment accuracy without any difficulty. It is straight forward and compatible to the conventional alignment methods. It also successfully demonstrated one of the most important advantages of DSE based alignment scheme: the alignment target can be totally random or regular and in micro-scale accuracy to achieve nanoscale accuracy. Armed with these understanding, we decide to focus our research and experiment in the true DSE based alignment procedure: IDMA.

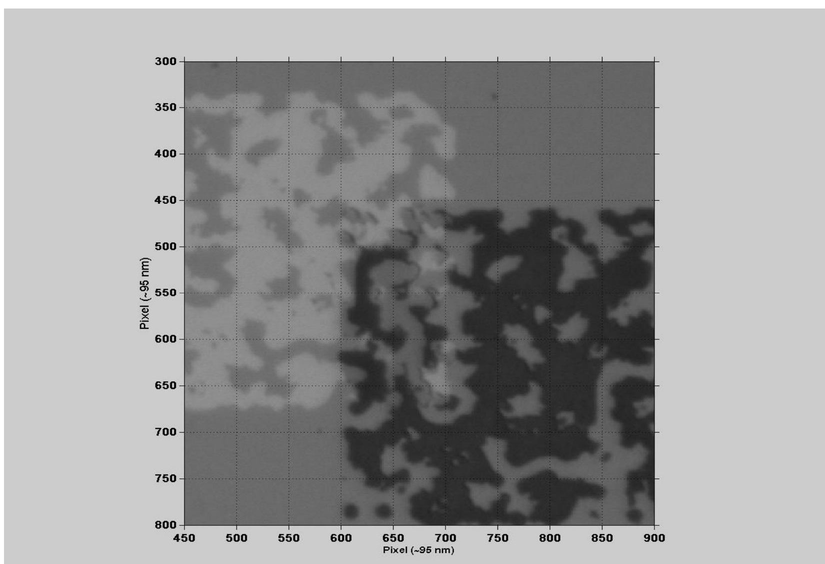

Fig. 8: DDMA alignment targets on substrate (lighter) and mold (darker)

4.2. IDMA, based on true displacement estimation

Indirect Displacement-Measurement-based Alignment (IDMA) entails acquiring alignment images from all substrates simultaneously in single unified images, to create reference images much like that depicted by Fig. 6a. When any one substrate moves with respect to the imaging system, nDSE algorithms can track its movement. Since all substrate marks are referenced off the same imaging system, changes to relative layer-to-layer positioning can be discerned by subtracting the individual substrate trajectories. The concept is illustrated in Fig. 6b, where the change in relative positioning of two separate features is determined through vector subtraction.

In the purest form of IDMA, our aim is simply to guarantee the consistent co-placement of materials. If we take a single snapshot of all the layers together, ensuring that we capture at least one mark from each layer in reasonable focus, and then move the materials about, we can use the snapshot as a reference to return the materials to their original relative locations, or a known offset to that location, if that is desirable. The advantage of the method is that we do not need to compare images across material boundaries. Each mark is correlated against a stored image of itself. We perform no pseudo-DSE measurements. All DSE measurements are true displacement measurements. Further, the marks need not be specially-prepared alignment marks. They could even be comprised of convenient device features.

Of course, all we can do with the procedural variant just described is to ensure repeatable co-placements. We are not directly measuring any sort of absolute layer-to-layer alignment. But, if we can perform an appropriate processing operation between the layers, such as a lithography operation, and then determine the absolute layer-to-layer alignments using a post-process inspection tool such as an SEM, we can then ascertain what layer-specific offsets would be required relative to the reference configuration to achieve the desired layer-to-layer co-placement (i.e. the alignment). With a reversible processing operation we could simply return the materials to their pre-processed state, and then, guided by DSE, position each layer to the desired offset off their positions in the reference image. Repeating the processing operation should now result in an accurate alignment.

Naturally, many processes are not reversible, and even if they were, we could not afford to double-process every wafer in this manor. However, if we could rely on run-to-run similarities (NOT layer-to-layer, but run-to-run), then we would need to perform this calibration sequence only once. With the stored reference image and the calculated offsets, we would achieve accurate and precise layer-to-layer alignment, run after run.

4.3. Practical hybrid methods based on both true- and pseudo-displacement estimation

An IDMA-only solution may define the most accurate method, but it may prove to be a bit impractical. With no features providing approximated absolute alignment, a human operator may find the method awkward.

The methods most likely to succeed will include IDMA and DDMA elements, with DDMA providing pseudo-displacement-based best-guesses, and IDMA providing the final word after receiving input from post-process inspections. An alternative, of course, is to use traditional geometric alignment marks (e.g. crosses and boxes) to provide the initial best-guess, and then to rely on IDMA from that point on. We are also studying methods which should theoretically be more tolerant of run-to-run variations in the marks. Going into greater development of these ideas is beyond the scope of this paper.

4.4. Planned experiments

We have designed nanoimprint molds for exploring IDMA concepts. We will use these molds on our-house nanoimprinter. The molds contain patterns which provide wafer and mold patterns for both IDMA and DDMA uses. The goal will be to provide overlay alignment metrology to better than 10nm. We will provide further details when we have completed the experiments

5. Advantages over current alignment methods

As stated in 3.2, lithographers traditionally judge mask-to-wafer alignment by overlaying complementary alignment patterns over one another until they appear to line-up in the field-of-view of an optical microscope. Geometric figures are used to take advantage of their known shapes, so that edge-detection and pattern matching techniques can help make-up for process-related variations, in particular across materials (e.g. between wafer and mask). Alternatives, which might include correlation-based methods similar to ours, do not provide the necessary tolerance to such process variations. Methods favored by nanoimprint lithographers include the use of moiré patterns, which also involve the creation of highly deterministic patterns, capable of extending the precision to the nanoscale, beyond the precision achievable through the use of typical geometric figures. We offer an alternative to the moiré patterns, relying instead on micron-scale alignment marks that need not be constructed to high precision. Our advantage is largely one of convenience and cost.

Since our methods rely on grayscale pixel correlations, we can use nearly arbitrary patterns as our alignment targets. The spatial spectral content of the targets can be broad and hence contain a great deal of information, an advantage over those methods requiring spectrally-pure targets such as the gratings required for producing moiré patterns. Most significantly, however, IDMA does not require the comparison of targets across material boundaries. In other words, marks on the mask or mold do not have to be compared against marks on the wafer, except to provide an initial estimate. Final precision is obtained by comparing mask marks to stored images of mask marks, and wafer marks to stored images of wafer marks. The measurements are therefore more precise, not susceptible to the process variations which distort the marks in random ways and require pattern-matching and edge-detection to recover the necessary precision.

The key is in providing reference images to displacement sensing, a concept introduced in 3.2, and developed in Section 4. We summarize the advantages of DSE-based alignment in terms of cost and simplicity as follows. The methods:

1. Use conventional optical systems to achieve nanoscale displacement sensing and overlay alignment.
2. Do not require specially-made high-accuracy alignment marks or displacement patterns.
3. Enable a potentially simpler and more robust system designs, and
4. Support flexible system architecture because the overlay alignment algorithms are implemented in algorithm and software.

6. Conclusion: enabling technology for nano- and emerging decentralized fabrication

Nano-imprinting lithography and a growing number of emerging fabrication technologies are posed to supplement and possibly eventually replace conventional centralized photolithography. Photolithography, in our opinion, is growing towards its limit, both because of its approaching physical limits and economically forbiddingly expensive. In the mean time, new and emerging applications in the area of bio-science, new sensor technologies, emerging display and consumer electronic industry, as well as post 9-11 military and homeland security applications demand diversified decentralized fabrication technologies. The era of one chip architecture fit all computers and a handful of chipsets determine the whole high tech industry is gradually being supplemented and replaced by on-demand, specialized, small- to medium-scale fabrication alternatives, such as nano-imprint lithography, step-and-repeat small processes, flexible substrate embossing/imprinting lithography, printed electronics, etc.

Meanwhile, the alignment requirements are also diversifying. Some fabrication techniques require less accurate alignment; such as laser lithography and roll-to-roll embossing for flexible displays, while others, notably nano-imprint lithography, require even higher alignment accuracy than conventional photolithography. Yet some new and emerging fabrication methods present new alignment challenges. For example, on flexible substrates, such as plastic or thin-film, the substrate under process is constantly changing its dimensions due to stress, thermo, and other influences. Conventional alignment methods would have difficulties performing in these new environments because the traditional alignment mark would deteriorate greatly under these conditions.

We clearly see this trend forward for the whole fabrication industry and decide to stay in front of one of the key area of new and emerging decentralized fabrication: alignment. The new and emerging decentralized fabrications require versatile and cost effective alignment methodology. Both in extreme high accuracy area of nano-fabrication and in demanding flexible substrate fabrication and printed electronics, the alignment methods have to be flexible, extendable, accurate, cost effective and in some cases better than start-of-the-art photolithography alignment requirements. We believe our DSE based alignment framework, based solidly on our 1 nm displacement sensing and estimation capability and our conventional optical plus algorithm implementation as well as real time, image feedback and flexible alignment mark requirement, makes a perfect alignment solution for several applications. We are very excited to carry out the experiments to demonstrate DDMA/IDMA and a number of hybrid alignment methods in nano-imprinting lithography to provide sub 10 nm alignments. We also investigating a wide array of applications, from printed electronic to focused ion beam milling to flexible substrate embossing and processing. Our DSE based alignment will play a center role in the alignment of all these applications.

REFERENCE

1. Kuglin C. D., Hines D. D., The Phase Correlation Image Alignment Method. Proceedings of the IEEE 1975 International Conference on Cybernetics and Society, pp. 163 - 165, September 1975.
2. Pearson J. J., Hines D. C., Golosman S., and Kuglin C. D., Video Rate Image Correlation Processor. SPIE Proceedings of the International Optical Computing Conference, San Diego, CA USA, pp 197-205, August, 1977.
3. Brown, G. Lisa. A Survey of Image Registration Techniques. ACM Computing Surveys. Vol. 24, No. 4, pp. 325-376, December 1992.
4. Foroosh (Shekarforoush) Hassan, Zerubia, B. Josiane. Extension of Phase Correlation to Subpixel Registration. IEEE Transaction on Image Processing, Vol., 11, No. 3, pp. 188-200, March 2002.
5. Shekarforoush Hassan, Berthod M., and Zerubia B. Josiane. Subpixel Image Registration by Estimating the Polyphase Decomposition of Cross Power Spectrum. INRIA, France. Res. Rep. 2707, 1995.
6. Stone Harold, Orchard Michael, Chang Ee-chien. Subpixel Registration of Images. Conference Record of the Thirty-Third Asilomar Conference on Signals, Systems, and Computers (Cat. No.CH37020). IEEE, Piscataway, NJ, USA. pp. 1446-52 vol.2, 1999.
7. Jiang J. Y., Hu X. D., Xu K. etc., Phase Correlation-based Matching Method with Sub-pixel Accuracy for Translated and Rotated Images. ICSP'02 Proceedings, 2002 6th International Conference on Signal Processing Proceedings, IEEE, pp. 752-5 vol.1, 2002.
8. Castro E. DE, Morandi C. Registration of Translated and Rotated Images Using Finite Fourier Transforms. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. PAMI-9 No. 5 pp. 700-703, September 1987.
9. Hoge W. S., Mitsouras D., Rybicki F. J. etc., Registration of Multi-dimensional Image Data Via Sub-pixel Resolution Phase Correlation. Proceedings 2003 International Conference on Image Processing, IEEE pp. II-707-10 vol.3; 2003.
10. Starikov Alexander. Metrology of Image Placement. Characterization and Metrology for ULSI Technology, 1998 Conference, AIP Conference Proceedings 449, pp 513-539, 1998.
11. Picciotto C., Gao J., Hoarau E., Wu Wei, Image Displacement Sensing (NDSE) for Achieving Overlay Alignment, Applied Physics A, Materials Science & Processing, Vol. A80, No. 6, page 1287-1299, March 2005
12. J. Gao, C. Picciotto, and W. Jackson: Displacement sensing and estimation theory and application, Applied Physics A, Materials Science & Processing, Vol. A80, No. 6, page 1265-1278, March 2005
13. Picciotto C., Gao J., Hoarau E., Wu W., Jackson W., and Tong W. M., Overlay alignment using optical microscopy and arbitrary surface features, Journal of Vacuum Science Technology B 23, 6, 3047-3051, Nov/Dec 2005