A new general approach for solving the self-calibration problem on large area 2D ultra-precision coordinate measurement machines

Peter Ekberg¹,², Lars Stiblert¹ and Lars Mattsson²

¹ Micronic Mydata AB, Box 3141, SE-183 03 Täby, Sweden
² Industrial Metrology and Optics, Production Engineering, KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden
E-mail: plekberg@kth.se

Received 12 April 2013, revised 27 November 2013
Accepted for publication 10 January 2014
Published 14 March 2014

Abstract
The manufacturing of flat panel displays requires a number of photomasks for the placement of pixel patterns and supporting transistor arrays. For large area photomasks, dedicated ultra-precision writers have been developed for the production of these chromium patterns on glass or quartz plates. The dimensional tolerances in X and Y for absolute pattern placement on these plates, with areas measured in square meters, are in the range of 200–300 nm (3σ). To verify these photomasks, 2D ultra-precision coordinate measurement machines are used having even tighter tolerance requirements. This paper will present how the world standard metrology tool used for verifying large masks, the Micronic Mydata MMS15000, is calibrated without any other references than the wavelength of the interferometers in an extremely well-controlled temperature environment. This process is called self-calibration and is the only way to calibrate the metrology tool, as no square-meter-sized large area 2D traceable artifact is available. The only parameter that cannot be found using self-calibration is the absolute length scale. To make the MMS15000 traceable, a 1D reference rod, calibrated at a national metrology lab, is used. The reference plates used in the calibration of the MMS15000 may have sizes up to 1 m² and a weight of 50 kg. Therefore, standard methods for self-calibration on a small scale with exact placements cannot be used in the large area case. A new, more general method had to be developed for the purpose of calibrating the MMS15000. Using this method, it is possible to calibrate the measurement tool down to an uncertainty level of <90 nm (3σ) over an area of (0.8 × 0.8) m². The method used, which is based on the concept of iteration, does not introduce any more noise than the random noise introduced by the measurements, resulting in the lowest possible noise level that can be achieved by any self-calibration method.

Keywords: self-calibration, algorithm, metrology, large area, ultra-precision, mask, CMM

Introduction
In the manufacturing process of large flat panel displays, square-meter-sized lithography photomasks are used. These masks consist of a chromium pattern on centimeter-thick glass or quartz plates. The pattern, typically made up of micrometer-sized lines and pixels, is written into a photosensitive layer on top of the chromium layer by a mask writer, also referred to as a pattern generator. The geometric demands on the X, Y-placement of the pattern features are exceedingly tight. The required capability is assured by using a chromium patterned reference standard, called a golden plate (GP), with known...
placement errors relative to an absolute Cartesian grid. This GP is used to calibrate the mask writer and the mask measurement tool in order to let the end user make use of photomasks manufactured by different pattern generators.

In previous papers, the principles of the mask writers and the performance of the mask metrology tool MMS15000 have been discussed [1–3]. The MMS15000 is today the world standard metrology tool for measurements of photomasks with dimensions of up to (1.3 × 1.5) m² [4]. The most important conclusion from previous papers [1, 2] was that the repeatability of the MMS15000, expressed as the standard deviation σ in the measured X, Y positions of hundreds of cross marks on a GP, is about 3 nm over a time frame of several hours. That level of stability of the measurement tool is an absolute requirement for being able to perform self-calibration of the entire system to a level where built-in deformations can be determined and finally be compensated for. In this paper, we present how we manage to calibrate the measurement tool to a perfect Cartesian grid using a novel self-calibration procedure.

As already pointed out, the most important prerequisite for achieving high absolute accuracy in X, Y measurements is that the measurement system is stable over time and shows good repeatability. In this context, it should also be explained that the display and mask manufacturing industries prefer to use the term registration instead of dimensional accuracy. Thus, in their terminology, the absolute registration requirement placed on the photomask pattern today is in the range of 200–300 nm (3σ) over an area of square meters. It is obvious that the calibration and verification of these numbers requires a 2D measurement machine with extreme stability and ultra-high accuracy. Besides local measurements of critical dimensions of features, e.g. line widths, the MMS15000 is capable of absolute registration measurements with a traceable accuracy better than 90 nm (3σ). To achieve this accuracy, a new self-calibration algorithm had to be developed, as there are no square-meter-sized traceable reference standards available today. The purpose of using the new algorithm is to find the absolute geometrical shape of both the stage of the measurement tool and the GP used in the calibration [3]. The self-calibration method presented in this paper differs in a number of aspects from the methods developed by Raugh and others in the 1990s [5–7]. Raugh put a lot of effort into deriving the mathematics behind self-calibration using a square lattice, and he also suggested implementations of the algorithms.

The fundamental principle of self-calibration is to solve an inverse problem and this is the same regardless of what kind of algorithm is used. It is in the details, generality and robustness against error sources that the differences between methods appear. With a measurement tool and a rigid plate filled with measurement marks, not necessarily with known X, Y locations, it is possible to achieve an absolute fit to a Cartesian coordinate system with a precision limited only by the noise level of the measurement tool. It is here assumed that the plate is rigid in the sense that it does not change shape when it is resting at different locations on the stage.

To minimize the noise level, a special measurement technique is used in the MMS15000. This technique does not rely on images grabbed by an ordinary CCD camera. Instead, we use our own developed measurement technique based on the principle of random sampling of pattern edges by a scanning laser beam. The measurement repeatability using this technique is better than 10 nm (3σ) when measuring a matrix of cells on an artifact of (0.8 × 0.8) m² [1, 2].

In figure 1, we show typical errors of a 2D measurement tool that can be assessed and compensated for by a self-calibration process. These include first, second and higher orders of deviations from a Cartesian grid. A typical first-order deviation example is the orthogonality, defined as the angle between the X-axis and Y-axis from 90°. Another example is the stage bow, defined as the deviation from a straight line of the X- and Y-axes, respectively.

Although we refer to a rigid body GP in this paper, it must be realized that an almost square-meter-sized plate cannot be generally considered as rigid when we speak of a few hundred nm absolute placement accuracy. Therefore, the self-calibration algorithm must be robust not only against ordinary measurement noise but also for the ‘shape instability’ caused by the plate itself when subjected to gravitational forces on a non-perfectly flat stage. We have implemented a special function called Z-correction that automatically corrects for these kinds of deformations as described in [8]. The principle used is to put the plate at several different stage positions in order to minimize the errors caused by the fact that the plate is not rigid, and also to generate enough non-redundant measurements. A third reason for using several positions on the stage is that there are no available quartz glass plates large enough to cover the entire measurement stage.

Before describing the theory of the self-calibration procedure, the definitions of the parameters are presented in the following section. Then, simulation results are presented and discussed, and finally experimental results are shown.

Definitions

Throughout this paper, we use the definitions given below. Some are also graphically represented in figure 2.

Grid. A matrix of points in the X, Y plane separated by well-defined distances, the so-called pitch. We use two different grids, each with its own pitch, in the self-calibration...
of the MMS15000. One is the grid of physical measurement marks on a plate and the other is the correction grid used for correction of the MMS15000 stage after the calibration.

\[ i_s, j_s \] are integers indexing points in the X and Y directions, respectively, when addressing plate points (marks) distributed over the plate (P) and points in a measurement matrix (M).

\[ (i_s = 0, 1, 2, \ldots, n_x - 1) \text{ and } (j_s = 0, 1, 2, \ldots, n_y - 1) \]

where \( n_x, n_y \) are the number of measurement marks on the plate in the X and Y directions, respectively. A nominal absolute position on the plate or stage is calculated by multiplying the index with the pitch in the corresponding direction, respectively.

\[ \text{Plate matrix (P) refers to the 8–16 mm thick large area plate with chromium measurement marks.} \]

On the nm scale, it is not perfect, and the ‘shape’ of the plate is described as a matrix of vectors \( \mathbf{P}(i_s, j_s) \) with their origin at the center of the plate.

\[ \text{Stage matrix (S). In analogy with the plate matrix, the 2D 'shape' of the measurement machine stage is described as a matrix of vectors } \mathbf{S}(i_x, j_y) \text{ relative to the origin at the lower left corner of the stage.} \]

\[ \text{Measurement matrix (M). Refers to a measurement of all plate measurement marks at a certain stage placement and rotation of the plate in the X, Y plane. A measurement of the measurement mark } i_x, j_y \text{ on the plate is expressed as } \mathbf{M}_k(i_x, j_y) \text{ and represents the vector coordinates relative to the center of the stage.} \]

\[ \text{Here, } k \text{ is the index of the measurement view representing a specific rotation/placement of the plate on the stage.} \]

\[ \text{Simulated measurement (SM). In the algorithm, simulated measurements are calculated based on the temporary representation of } \mathbf{S}, \mathbf{P} \text{ and the placement and rotation, } k, \text{ of each measurement } \mathbf{M}. \]

A simulated measurement at a certain location of the plate is expressed as a vector \( \mathbf{SM}(i_x, j_y) \).

\[ \text{Placement (plk) is a vector of two components } x_k, y_k \text{ containing the absolute coordinates of the averaged center point of all measurement marks } \mathbf{M}_k \text{ in a measurement view } k \text{ on the stage.} \]

\[ \text{Placement angle } \alpha(k) \text{ is obtained by calculating the angle from four corner marks of the measurement } \mathbf{M}_k \text{ when it has been back-rotated } (0, \pi/2, \pi, 3\pi/2). \text{ This angle is measured relative to the X-axis.} \]

**The self-calibration procedure**

The goal of the self-calibration is to find the absolute geometrical shape of the stage of the measurement tool relative to a perfect Cartesian grid. As a bonus, we will also determine the shape deviations of the plate measurement marks. Thus, the shape of the plate, determined by the self-calibration, can then be used as an accurate reference, a so-called GP, for the calibration of other machines. In other words, the self-calibration procedure is the method to use for the simultaneous calibration of both the stage and the GP. All deviations from the Cartesian grid can be found in the self-calibration procedure except the absolute scale, which needs a traceable calibration standard, e.g., the certified length artifact we use from Mikes [9]. However, the use of interferometers in an exceedingly well-controlled environment, for the position read-out of the measurement head of the MMS15000, guarantees that errors in scale will be small before any traceable scale correction is done. The input data used in a self-calibration are obtained from several measurements of the same measurement marks on the plate when it is placed in different positions and rotations on the stage. We will refer to the position/rotation as a view of the plate. At least three different views are needed in the self-calibration. Normally, more views are used for redundancy reasons and also in order to achieve good stability in the calibration along the edges of the plate as well as along the edges of the stage. In the large area display industry, it is very common to present these kinds of data using grid graphs with an absolute macro-scale combined with a local relative scale around each measurement mark. In figure 3, the principle of such a graph is shown.

In figure 4, we show a highly exaggerated example of a deformed plate with measurement marks (+) put on a deformed stage of the measurement machine. This is the starting point for the self-calibration process.

Masks used in the display manufacturing process are rectangular. Therefore, the matrix of measurement marks will not be a square matrix and the pitch between the measurement marks does not need to be the same in the X and Y directions. Throughout the self-calibration procedure, the deviations from the Cartesian grid, caused by the deformed stage, are internally described as a 2D function dependent on the absolute X and Y coordinates. By putting the plate with its measurement marks at different places on the stage, we obtain measurements that are samples of this 2D stage deformation function. This is based on
Figure 3. To visualize the deviation of a measured point from an absolute grid point, we use a local coordinate system (blue dashed square) centered at each grid point and connecting lines between neighboring points. The local scale, shown by the blue arrow, is typically of the order of tens to hundreds of nanometers, while the absolute distance between the grid points is of the order of several centimeters.

Figure 4. A highly exaggerated representation of a plate with measurement marks (+) separated by a pitch of \( pp_x \) and \( pp_y \) put on a deformed measurement machine stage with a virtual grid having a pitch of \( sp_x \) and \( sp_y \). The dashed lines refer to the perfect Cartesian grid of the stage. The goal of the self-calibration procedure is to obtain the corrections needed to compensate for the deviations in each stage grid point, to straighten out the deformed grid into the dashed grid.

The assumption that the deformed plate with its measurement marks is rigid with a grid that is constant and independent of the position of the plate on the stage.

Precise alignment of square-meter-sized and heavy plates is very difficult. The traditional self-calibration methods [5–7] based on different views (i.e. very accurate placements by rotations and translations of the plate on the stage) cannot be fulfilled in our case. Therefore, we had to develop a self-calibration method that does not require precise alignment. A great advantage with this method is that we may use different placement schemes for different sizes of plates and stages and it provides access to the weighting of certain areas on the stage for a more accurate estimation of the stage shape.

On the plate, we use a set of the \( n_x \times n_y \) measurement marks distributed in a matrix. The 'exact' locations of these marks are unknown but still they can be used to make a two-dimensional sampling of the MMS15000 stage deformation. Note that we will only have a representation of the shape of the plate at each discrete point \( i_x, j_y \) of the matrix \( P \). The shape of the stage may be found at every accessible \( X \) and \( Y \) point. This is because the sampling of the stage can be done at any arbitrary point and not necessarily only in the matrix grid \( i_x, j_y \), representing the shape of the stage after the calibration. It is more a question of what spatial frequencies of errors we are interested in finding. In other words, the method we have developed is not limited to finding deviations on the stage with spatial frequencies lower than the Nyquist frequency set by the pitch of the marks on the plate [10]. Such a limitation is present in some other inverse self-calibration methods. In practice, several placements and rotations of the plate are used to cover the whole measurement machine stage that is normally much larger than the plate used for the calibration. Multiple placements will also generate redundancy and level out the small geometrical distortions generated by the plate itself when put at different positions and rotations on the stage. In the self-calibration process, the placement scheme for the different measurements of the plate on the stage is important. As mentioned before, three different placements of the plate are enough for a calibration provided that these placements are not redundant. This means that new spatial information on the plate and stage must be retrieved in each measurement. By representing each measurement of the plate as a set of equations, a large number of placements will generate a huge set of equations. For three placements, this system of equations will be over-determined and for this reason possible to solve. This problem is typically solved using a least-squares fitting model and so it has been used by other authors regarding self-calibration [11]. But the interesting thing is that the system can actually be degenerate, i.e. the placements do not necessarily add information to generate a unique solution. To avoid this problem, we have developed a dedicated software package to optimize the placement scheme for a given stage size in relation to the plate size. The details of that are beyond the scope of this paper.

Theory

Principle of the self-calibration algorithm

We start with an unknown plate matrix \( P \) and an unknown stage matrix \( S \). The goal of the exercise is to find the absolute shape deviation of \( P \) and \( S \) relative to a perfect Cartesian grid.

When \( S \) is known, we can subtract this systematic correction so that the stage will describe a Cartesian system only limited by the uncertainty caused by random noise and not by the shape deviations.

\( S \) will contain information about scale differences in the \( X \) and \( Y \) directions as well as orthogonality. By aligning \( S_y \) to
the Cartesian Y-axis, the $x$ components will carry information about the orthogonality error [3].

In the algorithm, we use two different coordinate systems, one for $P$ with an origin in the center of $P$ and one for $S$ with the origin in the lower left corner of the stage $S$ (see figure 4).

The over-determined problem gives us more data points than unknown parameters and we use a figure of merit similar to the least-squares criteria to determine when we cannot gain any more information from the measured data. The implementation of the self-calibration is performed by an iteration loop, see figure 5, for solving the inverse problem.

We can express the principles of the algorithm as follows. Let us assume that we have a flat rigid object, in our case a plate with measurement marks, and we put this object on a metrology tool stage for measurement. If this plate is strictly rigid, the measurement marks will be identically the same in the plate coordinate system. If the measured positions of the measurement marks relative to the center of the plate are identical when put at different locations on the tool stage, we can claim that the stage coordinate system of the measurement tool is consistent. On the contrary, if the measurement mark positions vary with the placement of the plate, the tool coordinate system is not consistent and needs to be adjusted. The different appearance of the plate in different views and placements, $k$, is therefore a measure of how much the tool coordinate system differs from the perfect Cartesian system in the certain placement of the plate. Using this difference for the compensation of the tool stage function is nothing more than a calibration of the tool stage, but with the important difference that we do not necessarily know anything about the absolute positions of the measurement marks, i.e. the ‘shape’ of the plate. Therefore, we approximate the measurement mark pattern on the plate by using the average shape of this pattern from the different views as the best estimation of the plate shape. By virtually replacing this new object on the tool stage in the same locations and rotations as in the first measurements, we can then update the tool stage coordinate system at these locations. If the plate shape in a second comparison is the same in all placements, we are through with the calibration, if not, the iterative procedure is repeated.

To explain this in more detail, we introduce a temporary stage correction function $SCF$ which is a 2D function describing the correction of the stage $S$ for the current iteration loop. This function is represented as a matrix with the pitch $sp_x$ and $sp_y$ and has the same size as $S$. This matrix is set to zero at the initiation of the calculation. $M$, $P$ and $SM$ are all matrices defined earlier and represent different states in the calibration.

Finding a new version of the plate.

\[
P_{\text{new}} = [R^* (M + SCF_{\text{old}} - pl)]
\]  

(1)

The brackets $\{\ldots\}$ represent in this step the average of $k$ different measurement views, where $1 < k \leq N$. Each measurement has thus been done with different placements and rotations of the plate on the stage. For each placement, we calculate a new version of $P_{\text{new}}(k)$. Every placement and rotation also generates a calculated location vector $pl(k)$ and an associated rotation angle $\alpha(k)$ in $S$ that is saved.

The rotation $\alpha(k)$ is obtained from $M(k)$ by calculating the angle from four corner measurement marks as shown in figure 6(a) for a $-\pi/2$ rotation measurement.

The properties $pl(k)$ and $\alpha(k)$ are saved and updated in each iteration in the loop. The rotation of the plate on the stage is a multiple of right angle rotations, i.e. $0$, $\pi/2$, $\pi$ and $3\pi/2$ with a small offset angle $\alpha$ due to a non-perfect alignment relative to the stage coordinate system.

Each point in $M(k)$ that corresponds to an absolute coordinate in the stage coordinate system is now adjusted
by an amount equal to the interpolated deviation in the current version of $\text{SCF}_{\text{old}}$. A standard bi-linear interpolation is performed in the stage grid when this adjustment is calculated. After this operation, the result is back-rotated ($R^{-1}$) with the placement rotation, i.e. $0$, $\pi/2$, $\pi$ and $3\pi/2$ plus $\alpha$ as shown in figure 6(b). This can be expressed as

$$P_{\text{new}}(k) = R^{-1}_{\alpha(k)}(M(k) + \text{SCF}_{\text{interpolated}}^{\text{old}}(k) - pl(k)), \quad (2)$$

where $R^{-1}_{\alpha(k)}$ expresses the back-rotation of the measurement $k$ with the rotation angle $\alpha(k)$ plus a multiple of $\pi/2$, and $\text{SCF}_{\text{interpolated}}^{\text{old}}(k)$ is a matrix of the same size as $M(k)$ containing interpolated deviations in the measured points in $M(k)$.

The result $P_{\text{new}}(k)$ is now aligned with the stage X-axis in the sense that the four corners do not have any rotation. This operation will therefore effectively lock each $P_{\text{new}}(k)$ rotation to be zero and therefore secure that the rotation of the final average plate $P_{\text{new}}$ is zero.

Finally, we calculate a new version of $P_{\text{new}}$ as the average of all $P_{\text{new}}(k)$ that may be expressed as

$$P_{\text{new}} = \frac{1}{N} \sum_{i=1}^{N} P_{\text{new}}(i). \quad (3)$$

Verify improvement by simulated measurements.

$$\text{SM}_{\text{new}} = R^{+}(P_{\text{new}}) + pl - \text{SCF}_{\text{old}}. \quad (4)$$

In this step, a simulated measurement is performed based on the current version of the calculated coordinates of the measurement marks on the plate, the $\text{SCF}$ and the placement parameters (including placement angles). In other words, we ask the question ‘what would a measurement be using current plate, stage and plate placements?’

$R^{+}$ corresponds to a forward-rotation of a multiple of right angle rotations, i.e. $0$, $\pi/2$, $\pi$ and $3\pi/2$ minus $\alpha(k)$.

$N$ simulated measurements are created using the placement coordinates and rotations as

$$\text{SM}_{\text{new}}(k) = R^{+}_{\alpha(k)}(P_{\text{new}}) + pl(k) - \text{SCF}_{\text{interpolated}}^{\text{old}}(k), \quad (5)$$

where $R^{+}_{\alpha(k)}(P_{\text{new}})$ expresses the forward-rotation of the $P_{\text{new}}$ with the angle $\alpha(k)$, and then placed in the stage coordinate system $S$ by using the previously calculated placement $pl(k)$.

Update the stage correction function.

$$C = \text{SM}_{\text{new}} - M. \quad (6)$$

In this step, we calculate the correction $C$ based on the simulated measurements and the real measurements in the stage coordinate system. This correction is then used to update the $\text{SCF}$ before the next iteration. Obviously, $C = 0$ if we have managed to calculate an absolutely correct $\text{SCF}$ and $P$. So by calculating the standard deviation of the elements of matrix $C$, we obtain a figure of merit of the calibration. The value of this standard deviation is one of the termination criteria we use in the loop.

As the stage grid and plate measurement marks do not normally coincide, the simulated measurements $\text{SM}_{\text{new}}(k)$ and the real measurements $M(k)$ will be located somewhere between the stage grid points. Therefore, a special ‘backward’ bi-linear interpolation method is used to distribute the difference $\text{SM}_{\text{new}}(k) - M(k)$ to the elements in the $\text{SCF}$.

We introduce the vector $Ce(x_i', y_j')$ as

$$Ce(x_i', y_j') = \text{SM}_{\text{new}}(k) (i, j) - M(k) (i, j). \quad (7)$$

Here, $x_i', y_j'$ represents the coordinates, in the stage grid, of a plate measurement mark placement. The measurement mark will thus normally be surrounded by four approximately correct stage grid points given by the $\text{SCF}$. The coordinates $x_i', y_j'$ are extracted from the placement matrix $pl(k)$ and $\alpha(k)$.

In figure 7, the principles of this method are shown. Only the nearest neighbors of the elements in the $\text{SCF}$ to the measured $Ce(x_i', y_j')$ point will be adjusted in proportion to how far they are from that point. In the graph, this is exemplified for the $\text{SCF}(i_{sx}, j_{sy})$ with the distances ($dx$, $dy$).

The position of $Ce(x_i', y_j')$ is thus a measure of how much of the stage $\text{SCF}$ shall be updated in the surrounding points. The update is only performed at grid points inside the rectangle defined by $x_{filter}$, $y_{filter}$ from the point $x_i', y_j'$ in the correction map. $x_{filter}$ is set to be equal to $pp_x$ and $y_{filter}$ is set to be equal to $pp_y$. These settings will guarantee that a point $x_i', y_j'$ will update all surrounding grid points in $\text{SCF}$ inside the filter and not leave any non-updated points between the measured points in $M(k)$.

The distribution of the adjustments to the grid points in the rectangle is performed by backward bi-linear interpolation based on distances between $Ce(x_i', y_j')$ and the grid points. Referring to figure 7, the correction $\text{SCF}(i_{sx}, j_{sy})$ is calculated from $Ce(x_i', y_j')$ as

$$c(i_{sx}, j_{sy}) = |x_{filter} - dx|/x_{filter} \cdot Ce(x_i', y_j') \text{ if } |dx| \leq x_{filter}$$

$$c(i_{sx}, j_{sy}) = 0 \text{ if } |dx| > x_{filter} \quad (8)$$

and

$$\text{SCF}(i_{sx}, j_{sy}) = \text{SCF}_{\text{old}}(i_{sx}, j_{sy}) + \text{SCF}_{\text{old}}(i_{sx}, j_{sy}) \text{ if } |dy| \leq y_{filter}$$

$$\text{SCF}(i_{sx}, j_{sy}) = \text{SCF}_{\text{old}}(i_{sx}, j_{sy}) \text{ if } |dy| > y_{filter}. \quad (9)$$

where $dx$ is the distance between the stage grid point $i_{sx}$ and the point $x_i'$ and $dy$ is the distance between the stage grid point $j_{sy}$ and the point $y_j'$. $C_{\text{old}}(i_{sx}, j_{sy})$ is the current value in cell $i_{sx}, j_{sy}$ that has been updated by other points in $M$.\n
\begin{figure}
\centering
\includegraphics[width=0.8\textwidth]{figure7}
\caption{The red point $Ce(x_i', y_j')$ is a plate measurement mark that is placed somewhere between four approximately correct stage grid points (black) represented by the $\text{SCF}$, which is in the process of being updated.}
\end{figure}
We start with an approximately correct but at the nm scale unknown stage and plate shape. As a first approximation, we assume that the stage is described as a perfect Cartesian grid, i.e. the SCF is just a table of zeros. If this assumption is correct, we must measure the same shape of the plate regardless of where it is placed on the stage. If this assumption is wrong, the best estimate we can make of the plate shape is to use the average shape of the plate as calculated from the measured measurement marks but represented in the plate coordinate system. After this correction of the marks on the plate, it is replaced on the stage in the same positions and rotations as before. After this replacement, we simulate a series of measurements. The difference between these new simulated measurements and the real measurements is a measure of how much we must modify the SCF. After this modification, we generate new measurements, using the current version of the SCF, that now differs less than the first one, where we assumed that the stage was perfect. By repeating this iterative procedure, we will gradually enhance the plate shape and the SCF. After the minimum error criteria have been reached, the process is terminated. As long we have measurements that are independent, i.e. no degeneration exists, this procedure will converge to the best estimation of both the shape of the plate and the SCF. In the following section, we will show the performance that can be achieved in large area measurements using this self-calibration approach.

Results and discussion

In order to demonstrate the performance of the self-calibration algorithm, we will perform two simulations and one case study of real data. The first simulation will illustrate the self-calibration of a \((0.75 \times 0.80)\) m\(^2\) area of the measurement tool stage function simultaneously with a deformed plate grid having 272 measurement marks. In the second simulation, the algorithm is subjected to random noise to test its robustness against variation in input data. As a final test of the algorithm, real measurement data are used from 380 measurement marks covering an area of \((0.798 \times 0.756)\) m\(^2\). In this case, the result is demonstrated as an overlay, i.e. how well the measurement marks of four independent measurements obtained at different rotations of the plate deviate from the average location of the four.

To avoid degeneration in the calibration process, the placement scheme of the plate is vital. For example, to extract orthogonality, rotations of \(\pi/2\) must be involved for the plate. The algorithm we have developed therefore calculates a so-called coverage map for estimating the number of measured placements and rotated measurements covering a certain area around each grid point on the stage. Thus, a warning is raised if a certain placement scheme is degenerate or suffers from bad coverage. A detailed description of this is out of the scope of this paper.

**Simulation of plate and stage shape**

In the first simulation, we demonstrate a simple case of using four placements of a plate with \((16 \times 17)\) measurement marks having a pitch of 50 mm in both \(pp_x\) and \(pp_y\). The slightly deformed shape of this plate is shown in figure 8(a) using the multi-scale presentation introduced in figure 3.

For the simulated two-dimensional stage function, we use a grid of \((30 \times 30)\) points with a pitch of \((40, 40)\) mm in the \(X\) and \(Y\) directions, respectively. We simulate a typical bow of the stage of about 5 \(\mu\)m in the \(Y\) direction as represented in the figure 8(b).

The first plate placement we use in this simulated measurement is the upper left corner of the stage. We aim the corner of the plate close to the absolute position \((100,150)\) mm in the \(X\) and \(Y\) directions, respectively. Then we rotate and place the plate according to table 1 relative to the first plate position. The overlapping plate grid positions for the four cases are illustrated in figure 9.

By performing 10 000 iterations of the calibration loop shown in figure 5 and using the data obtained from the four placements/rotations in table 1, we obtain a result confirming the convergence of the algorithm. The result is presented as a grid difference graph between the simulated input plate (figure 8(a)) and the calculated shape of the output plate and is a measure of the precision of the calibration. Another measure is the difference between the simulated stage grid and the final calculated shape of the stage grid function as delivered by the algorithm. In the latter case, it is only relevant to compare the output stage in the area covered by the placements of the plate. These two comparisons are shown in figures 10(a) and (b), respectively, showing that we have calculated the shape very accurately of both the plate and the stage. Note that the local scale revealing the errors is now reduced to 20 nm.

The \(3\sigma\) variation for the difference of the plate is 3.7 nm and 2.6 nm in the \(X\) and \(Y\) directions, respectively. The \(3\sigma\) variation for the difference of the stage is 4.5 nm and 3.3 nm in the \(X\) and \(Y\) directions, respectively.

In the calibrated area, it is likely that we have small absolute scale errors of the stage if it is compared to a traceable meter standard with higher accuracy than the interferometer-based systems we use in stage movements. For the same reason, the plate may have an absolute scale error. The algorithm will automatically move all scale errors into the matrix \(P\) that is a description of the plate. Since absolute scale needs to be calibrated using a traceable reference, the effect of the scale has been removed in the graphs in figure 10.

As can be seen, the deviations along the edges are worse compared to the middle of the calibrated area. The reason
Figure 8. The \((0.75 \times 0.80)\) m\(^2\) plate grid (a) and the \((1.2 \times 1.2)\) m\(^2\) large stage grid (b) used in the simulation. Note the dual scales used in each graph as described in figure 3.

Figure 9. The four different simulated placements of the \((0.75 \times 0.80)\) m\(^2\) large plate with the 50 mm plate grid shown in red, green, pink and orange on top of the blue 40 mm stage pitch grid. The local scale around a grid point (cf figure 3) is in this case 3 mm. The background grid (dashed in the figure) has a pitch of 50 mm in both \(X\) and \(Y\).

for this difference is the limited coverage of grid points at the edges, three measurements, compared to four different measurements in the middle.

Simulation of robustness against noise

In the second simulation, we add some random noise to the measurement. The three sigma variation of this noise is 18 nm in the \(X\) and \(Y\) directions, respectively. In this test, we will avoid interpolation errors and therefore we use the same pitch, 50 mm, of the plate and stage grids in the \(X\) and \(Y\) directions, respectively. Table 2 shows the placement scheme.

The result after self-calibration yields a 3\(\sigma\) variation for the plate of 18 nm and 17 nm in the \(X\) and \(Y\) directions, respectively, and the corresponding 3\(\sigma\) variation for the stage is 16 nm and 15 nm, respectively, in the \(X\) and \(Y\) directions.

Table 2. The different rotations and translations used in the simulation with noise added. The result of this test is presented in figure 11.

| Placement | Rotation angle | Translation \(X\) | Translation \(Y\) |
|-----------|----------------|-------------------|-------------------|
| 1         | 0              | 0                 | 0                 |
| 2         | \(\pi\)        | 0                 | 50                |
| 3         | \(\pi/2\)      | 0                 | 0                 |
| 4         | 3\(\pi/2\)     | 0                 | 50                |

This simulation shows an ideal case never occurring in practice, but it is relevant for demonstrating the robustness of the algorithm against noise. As can be expected for a high performing system, the calculated output noise will be of the same order as the added input noise of the system.

Real measurement data

As a last demonstration of the performance of the algorithm, we will use real input data obtained from our MMS15000 measurement tool with the following placement scheme.

The pitch of the \((20 \times 19)\) matrix of measurement marks on the plate in this case is \((42, 42)\) mm in the \(X\) and \(Y\) directions, respectively. Since we do not know the exact shape of the reference plate used, we need another measure of the quality of the calibration. So instead of comparing the output plate with a simulated input plate, we will measure the overlay, i.e. showing how well the measurement marks of four independent measurements at different rotations of the plate deviate from the average location of the four after the calibration algorithm has been applied. Each placement of the GP according to table 3 will thus generate its own output plate. We can express a measurement using the principal equation

\[ M = P - S. \]
Figure 10. (a) The remaining differences between the grid shape of the calculated output plate and the simulated input plate grid (cf figure 8(a)) shown on a local scale of 20 nm around each grid point. (b) The remaining differences between the calculated output stage grid and the simulated stage grid in the covered area of the placements, plotted on the same local scale as in (a).

Figure 11. (a) The difference between the calculated output plate grid and the simulated plate grid when a random measurement noise of 18 nm (3σ) has been added. (b) The same difference as in (a) but for the calculated output stage grid and the simulated stage grid in the covered area of the placements.

Table 3. The four different placements of a reference plate.

| Placement | Rotation (radians) | Translation X (mm) | Translation Y (mm) |
|-----------|--------------------|--------------------|--------------------|
| 1         | 0                  | 0                  | 0                  |
| 2         | π                  | 0                  | -42                |
| 3         | π/2                | 42                 | -42                |
| 4         | 3π/2               | 0                  | -42                |

where $P$ is placed in $S$ as shown in table 3. In each placement, we therefore can calculate a plate as

$$P(i) = M(i) + S,$$

where $i$ is the index of the placement. If we then back-rotate each $P(i)$ (except the first one when $i = 1$) to 0°, we have four different versions of the plate corresponding to each measurement $M(i)$.

The output plate $P_{ave}$ is the average of these four calculated plates that may be expressed as

$$P_{ave} = \frac{1}{4} \sum_{i=1}^{4} P(i).$$

In figure 12, the average plate and the overlay of the four plates are presented. In the overlay plot, each plate is compared with the average plate $P_{ave}$.

In table 4, the 3σ variation of the difference between the measured marks for each placement and the average position of the marks is shown in the X and Y directions.

As seen in table 4, the deviations of the different plates compared to $P_{ave}$ are <150 nm (3σ). These are typical
Figure 12. The average plate $P_{\text{ave}}$ and the overlay between the different placements of the plate. The local scale around each grid point is 500 nm for $P_{\text{ave}}$ and 200 nm for the overlay.

Table 4. The deviation of each output plate compared to the average plate expressed as $3\sigma$.

| Placement | X $3\sigma$ (nm) | Y $3\sigma$ (nm) |
|-----------|------------------|------------------|
| 1         | 132              | 98               |
| 2         | 149              | 76               |
| 3         | 74               | 84               |
| 4         | 94               | 95               |

numbers in a self-calibration of the MMS15000 over an area of approximately $(0.8 \times 0.8)$ m$^2$. The registration specification of the MMS15000 is 90 nm ($3\sigma$) for a verification measurement of the placements $0$, $\pi/2$, $\pi$ and $3\pi/2$ inside the calibrated area. It is not possible to fulfil this specification without correcting for the fact that the plate is not rigid during the calibration. As mentioned before, a special function called Z-correction is used for this purpose [8]. By using this function in a self-calibration, we have achieved a verified uncertainty around 50 nm ($3\sigma$) over an area of $(0.8 \times 0.8)$ m$^2$. In summary, we have shown by several simulations and also by use of real data that the presented self-calibration algorithm and its iteration loop converge and give robust results even when subjected to noise disturbances and arbitrary placements. We cannot prove convergence in a strict mathematical sense, although it is possible to prove it if we apply the same constraints as Raugh [5] did, i.e. exact placements of the plate, pitches of measurement marks and pitches used on the stage.

Conclusions

We have shown that by using a new 2D self-calibration algorithm for large area measurement machines we can achieve a calibrated precision in the nm range over square meters that is limited only by the noise level of the measurement tool. This is achieved by an iteration loop resulting in the stepwise improvement of fitted functions describing the shape of both the plate with measurement marks and the shape of the measurement machine stage. By simulations, we have verified that as long the placement is not degenerate, the algorithm converges to a correct solution. The presented results show that it is possible to achieve an absolute $3\sigma$ uncertainty of 150 nm over $(0.8 \times 0.8)$ m$^2$ by comparing the different plates corresponding to the different placements on the stage of the MMS15000 measurement tool. The presented algorithm is general in the sense that fewer constraints are needed compared to other self-calibration algorithms. The quality of the calibration can be measured in a robust way both by using simulations and also by comparing the different views of the plate on the stage when real data are used. To further improve the calibration, it is necessary to compensate for the mechanical deformations of the plate when it is placed on the stage. How this is done is the subject for a follow-up paper on our developed Z-correction method.

It is possible to calibrate the entire measurement tool stage using a much smaller plate, either by first finding the shape of the plate with its measurement marks (in a smaller area of the stage) and then using it as a reference for a calibration of the larger area, or by designing a placement scheme covering the entire stage and in this way finding the shape of the stage and plate simultaneously. A full self-calibration plus verification of a large area stage $(1.5 \times 1.5)$ m$^2$ takes several days of time including tens of plate measurements. This means that an extremely time- and temperature-stable measurement machine has to be used if nanometer scale performance is to be achieved.

Acknowledgment

This work has been supported by Micronic Mydata, Stockholm, Sweden.

References

[1] Ekberg P, Stiblert L and Mattsson L 2010 Ultra precision geometrical measurement technique based on a statistical random phase clock combined with acoustic-optical deflection Meas. Sci. Technol. 21 125103
[2] Ekberg P, Stiblert L and Mattsson L 2012 Large area ultra-precision 2D geometrical measurement technique based on statistical random phase detection Meas. Sci. Technol. 23 035007

[3] Ekberg P 2013 Development of ultra-precision tools for metrology and lithography of large area photomasks and high definition displays PhD Thesis KTH Royal Institute of Technology, Stockholm

[4] Micronic Mydata AB MMS www.micronic.se

[5] Raugh M R 1997 Error estimation for lattice methods of stage self-calibration Proc. SPIE 3050 614

[6] Raugh M R 2002 Self-calibration of interferometer stages Technical Report ARITH-TR-02-01 http://arith.stanford.edu/tr/StanfordSelfCalReport-200308.pdf

[7] Ye J, Takac M, Berglund C N, Owen G and Pease R F 1997 Exact algorithm for self-calibration of two-dimensional precision metrology stages Precis. Eng. 20 16

[8] Ekberg P, Stiblert L and Mattsson L 2014 Z-correction, a method for achieving ultraprecision self-calibration on large area coordinates measurement machines for photomasks Meas. Sci. Technol. 25 055002

[9] MIKES A 750 mm long quartz artifact with certified expanded uncertainty (k = 2) of 116 nm www.mikes.fi Certificate of Calibration M-05L37 (Espoo: MIKES)

[10] Proakis J and Manolakis D 2007 Digital Signal Processing: Principles, Algorithms, and Applications 4th edn (Upper Saddle River, NJ: Pearson Prentice Hall)

[11] Raugh M R and Minor J M 1996 Statistical perspectives of self-calibration Proc. SPIE 2725 114