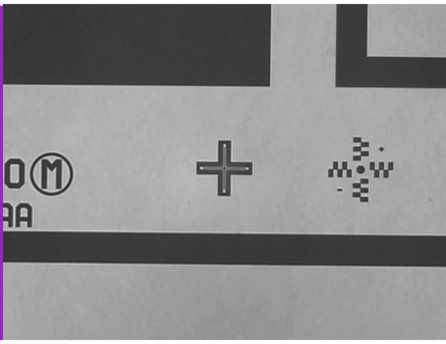


Profiler Pattern Recognition Selection

Tencor™ P-Series and HRP®-Series

Introduction

The stylus surface profiler, with its ability to precisely measure step heights and surface topography, is a key metrology tool for the development and monitoring of various semiconductor and thin film head processes such as plating, film deposition, etch, metal and oxide Chemical-Mechanical Planarization (CMP), and backside grind. Not only has the surface profiler proven its capability as a development and monitoring tool, but the surface profiler meets the demand for automation and throughput as required by production environments.

Measurements for production and development fabs can be fully automated on the Tencor™ P-series and HRP-series stylus profilers. Automation allows engineers and operators to start a recipe either locally or through factory automation software, where sample load, pattern recognition, measurement, and analysis steps are all completed automatically. This automated metrology requires advanced calibrations for measurement precision and tool matching, automated cursor placement, and robust pattern recognition. This application note describes pattern recognition and its usage for automated sample alignment for the Tencor P-7, P-17, P-170, and HRP-260 profilometers.

Pattern Recognition Alignment

KLA Instruments™ stylus profilers with pattern recognition allow the user to setup sequence recipes that automatically align the sample, enabling automated measurements. Alignment marks, such as those used for stepper lithography processes, are taught first, with the subsequent measurement sites taught with coordinates relative to the alignment marks. When samples are measured, the same alignment marks are located and the sample position adjusted to remove any positioning error due to sample loading. This full automation removes the need for the operator to manually locate and measure scan sites.

The profiler software also supports manual alignment methods. If pattern recognition fails, the user can manually locate the feature to assist pattern recognition. Single Deskew uses a single, manually-located site, and the user can then correct one or more axes: X, Y, and/or Theta. No Deskew is also available for samples that do not have a pattern.

Pattern Recognition Deskew

Three modes of pattern recognition are available: (1) pattern recognition deskew, (2) deskew with site pattern recognition, and (3) site pattern recognition. Pattern recognition with deskew is the most commonly used method to align the sample. The user teaches two alignment marks, called deskew sites, and the software records the X, Y and Theta position of each site. When a sequence recipe is executed, the profiler software automatically moves to the pre-programmed X, Y, and Theta coordinates for each deskew site. By comparing the location of the taught deskew site to the recognized deskew site, the system can then correct for translational and rotational error. These corrections are applied to the scan sites, enabling automated measurements.

Pattern recognition deskew can be performed with single or double deskew. Single deskew finds both deskew patterns once and calculates the XY correction for each scan site to remove X, Y and Theta error. Single deskew is the most common method and is typically the fastest. Double deskew finds both deskew patterns twice. In the first pass, the Theta error is calculated, and a Theta rotation is performed. In the second pass, the X-Y error is calculated, and an X-Y translation is performed. Double deskew is used when there is large theta error in sample loading or the user needs very precise theta alignment for either measuring small features or accurate width calculations.

The location of pattern recognition deskew sites is crucial to alignment accuracy. The optimum distance from the edge of the sample is two-thirds of the substrate radius (e.g., 33mm from the edge of a 200mm wafer). The optimal locations are in opposite quadrants of the sample. Figure 1 shows two example positions for best deskew sites, either X_{a1} and X_{a2} or X_{b1} and X_{b2}.This positioning allows for the most accurate calculation of translational and rotational error. If deskew sites are positioned close to the center or along the same vertical or horizontal axis, the accuracy of the rotational correction will be reduced.

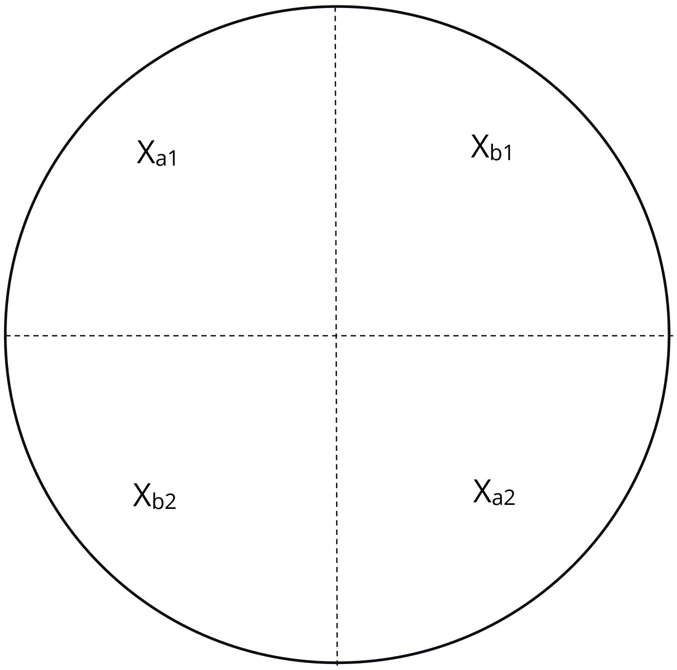


Figure 1. Wafer diagram showing two pairs of example positions for best deskew sites, either Xa1 and Xa2 or Xb1 and Xb2.

Site Pattern Recognition

Site pattern recognition performs pattern recognition on a feature that is close to the measurement site. The X and Y error correction is only applied to the measurement site(s) that are associated with the alignment feature. Site pattern recognition is often used if only one or two sites are being measured, since it will be faster than performing pattern recognition deskew.

Deskew with Site Pattern Recognition

Deskew with site pattern recognition performs the same deskew as described earlier, but with an additional pattern taught near each measurement site. This method is not typically used, due to the additional time required for pattern recognition and the relatively small reduction in site positional error. However, for certain applications, this method can be useful to reduce error introduced by transferring a recipe between two tools. In addition, if there are multiple scan sites in close proximity to the alignment site, the **Use Previous** option can be used in order to apply a single site pattern recognition error correction to multiple sites.

Types of Pattern Recognition

KLA Instruments stylus surface profilers offer four methods for pattern recognition matching: Contrast, Edge, Enhanced Edge and Geometric Model Find. Contrast and Edge are the oldest methods that are available, and Enhanced Edge and Geometric were introduced with the release of Windows 7 software. Geometric Model Find typically performs the best and is the most flexible, allowing for variation in lighting, scale (zoom) and rotational angle.

Contrast Matching

Contrast pattern recognition uses a normalized grayscale correlation to match features, requiring high contrast features to achieve the best performance. In addition, the contrast should remain relatively constant from sample to sample since it is matching the absolute grayscale value of the taught pattern recognition model to the sample.

Figure 2 shows a good pattern for a Contrast alignment mark. This pattern has a large difference in grayscale value between the pattern and the surrounding surface.

Figure 2. Example of a good pattern for Contrast pattern recognition

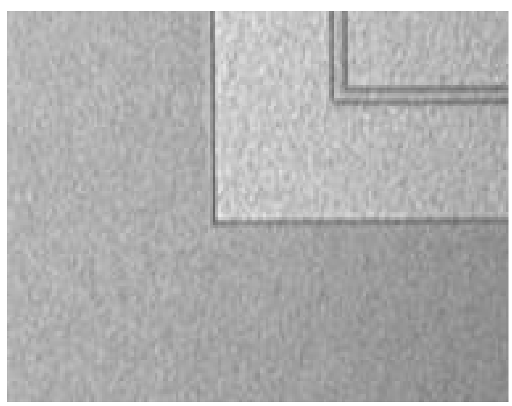


Figure 3. Example of a poor pattern for Contrast pattern recognition.

Figure 3 shows a poor pattern for a Contrast alignment mark, with little grayscale difference between pattern features. Applications such as polysilicon deposition or surfaces with small steps (< 50nm) and only one material type exposed give lower contrast patterns.

Edge Matching

Edge pattern recognition uses filters to smooth the model and enhance the edges of features. Pattern recognition is still based on a normalized grayscale correlation to match features, but it is relative rather than absolute. This method allows the grayscale values to change, but will still give good pattern recognition performance as long as feature edges are still present. Figure 4 shows a good pattern for an Edge alignment mark, with multiple horizontal and vertical edges.

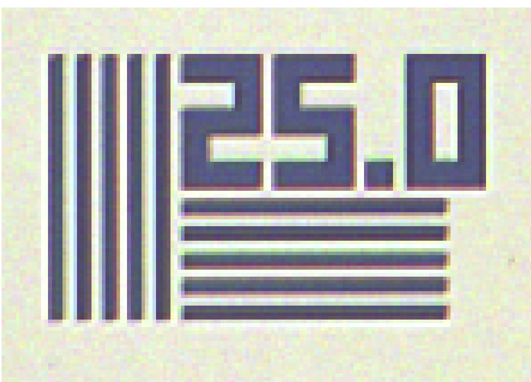


Figure 4. Example of a good pattern for the Edge pattern recognition

Figure 5 shows a shallow trench isolation (STI) CMP pattern recognition example. For CMP, typically the grayscale (contrast) will change as a function of one or more CMP process parameters, such as time or polishing amount, and can result in features flipping between dark and light. In Figure 5, the horizontal and vertical lines as well as the box features, change contrast, but remain in the same location. Edge pattern recognition performs well on these strong-edged patterns, since it uses relative changes in grayscale, rather than absolute.

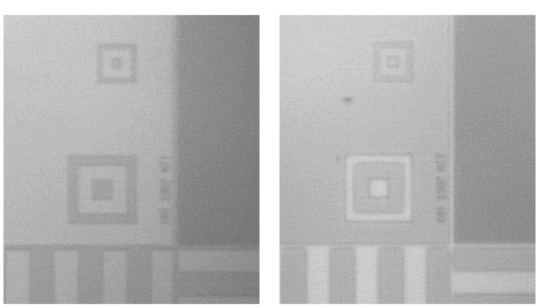


Figure 5. STI changes in contrast as a function of CMP conditions

Enhanced Edge Matching

Enhanced Edge pattern recognition uses the same matching method as Edge pattern recognition. The difference is the filters used, with Enhanced Edge resulting in less smoothing of the pattern recognition alignment model. For some patterns, Edge mode over-smooths and blurs the pattern line edge(s), degrading pattern recognition performance. In this case, Enhanced Edge maintains the sharp edge lines and provides a robust solution for pattern recognition.

Figure 6 shows an example where Enhanced Edge will perform better than Edge pattern recognition. This pattern has very thin and shallow edge lines, such that over-smoothing from Edge pattern recognition would blur the lines. The reduced filtering of Enhanced Edge works well, due to less feature smoothing.

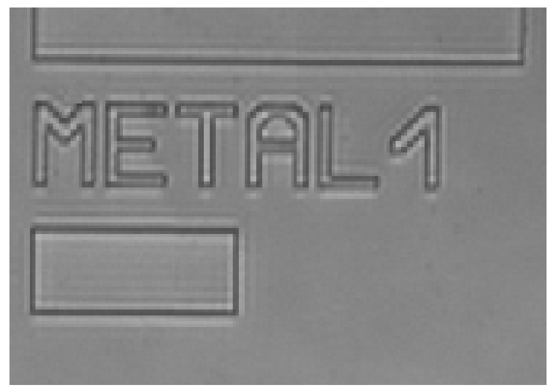


Figure 6. Example of a good pattern for Enhanced Edge pattern recognition

Geometric Matching

Geometric or Geometric Model Find pattern recognition adds shapes to improve matching accuracy. Adding the shape improves matching, as compared to using only edges and model grayscale values. Geometric generally provides the best performance, and it is the recommended starting point for a new sequence recipe.

Figure 7 shows an example of two different alignment models and where these features are found in the image. The L and T shapes are each uniquely identified, including shapes that have a large rotation relative to the models. Loaded samples typically have only a small rotational angle from handler or manual loading, so by default, the software search range is limited to ± 5° of rotation.

Figure 8 compares the performance of Edge versus Geometric for different zoom levels (different scales) using the optical zoom capability on the P-17 and P-170. The scale may have small variations due to optics alignment or camera pixel size, or large variations due to transferring recipes between tools with different optics designs. In addition, Geometric is less sensitive to light level variations, making Geometric pattern recognition the best method to overcome differences between tools when transferring recipes.

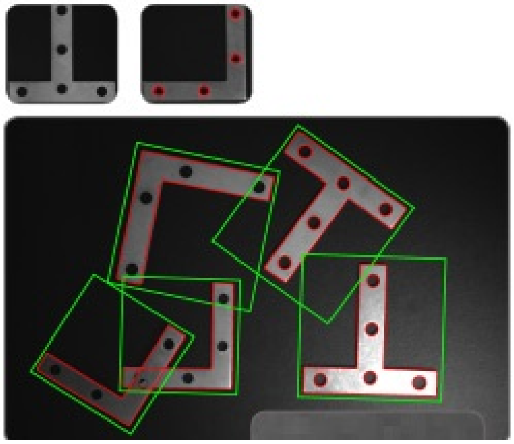


Figure 7. Example of Geometric model matching

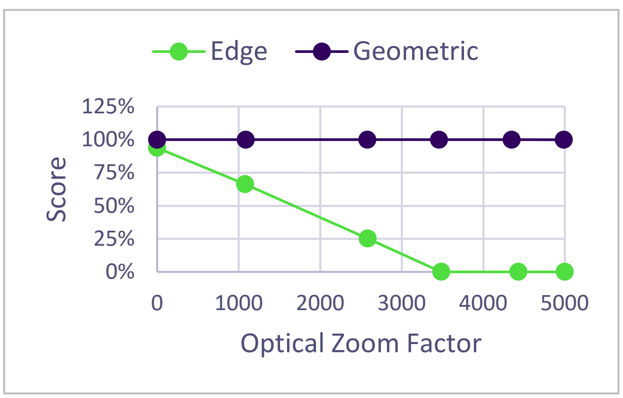


Figure 8. Edge versus Geometric match score as a function of zoom (scale)

Other Factors Impacting Pattern Recognition Performance

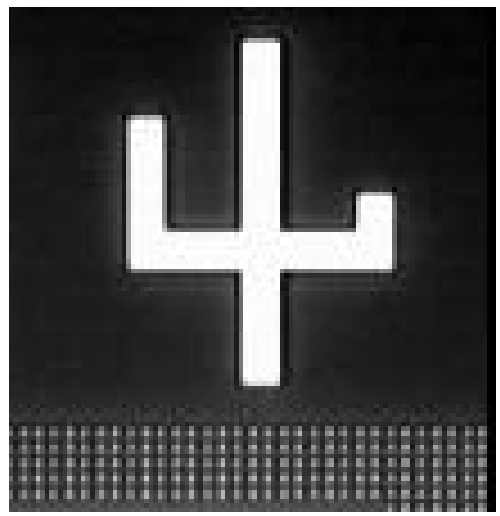
The feature shape, size, scale, and brightness/contrast can all impact pattern recognition performance. Typically the stepper model is used for pattern recognition, but it is not always the best model for good pattern recognition performance.

Pattern Recognition Alignment Models

Patterns with horizontal and vertical lines are best. Any circular pattern, such as the one shown in Figure 9, is a poor choice, but especially bad for Geometric since circular patterns do not include any information about the feature rotation.

Avoid patterns with small, intricate patterns. Figure 10 shows an example where the white pattern is very good for pattern recognition, but the small intricate pattern at the bottom includes one or two pixels with different contrast. After filtering, this pattern will be smoothed out and will not provide a unique feature for pattern recognition.

Figure 9. Circular patterns are rotationally symmetric, so are a poor choice for Geometric pattern recognition.



.

Figure 10. The intricate pattern at bottom of this image makes this image a poor pattern recognition model. However, if the lower pattern is not included with the image, the upper pattern will work well.

Image Brightness and Contrast

There are four different methods to set the image brightness and contrast: Default, Fixed, AutoLight, and Advanced AutoLight. Default uses the last set brightness and contrast, which is useful if all samples being measured can utilize the same settings. Fixed uses the brightness and contrast that were set when the sequence recipe pattern recognition model was taught. AutoLight adjusts the brightness and camera contrast until the best light level is achieved (this method is not available on the HRP-series profilers). Advanced AutoLight adjusts only the brightness until the best light level is achieved.

Fixed is the recommended method to set the light level, because the user has specified the best settings for the sample during teaching. In addition, Fixed ensures that the same settings are used for all samples. This method is the most robust, since there is no user error from using Default and there is no potential to fail to find a good light setting, which might be the case with AutoLight or Advanced AutoLight. The Fixed method is also the fastest since no time is spent adjusting the brightness and camera contrast. When combined with Geometric pattern recognition, Fixed provides good performance to account for process and tool-to-tool matching variation.

Image Scale

For best performance, any differences in scale should be minimized. The P-17 and P-170 optical zoom can be adjusted until the video calibrations are matched between tools. Figure 11 shows the impact on the pattern recognition score as the difference in scale increases for Edge pattern recognition. However, for systems like the P-7 and HRP-260 that have different optics, matching is not possible. For matching between systems with large changes in the optics, such as matching a P-170 to the HRP-260, Geometric can be used with appropriate scale search settings. However, the best performance is achieved when scales are matched, so for large changes, it is recommended to re-teach the pattern recognition model.

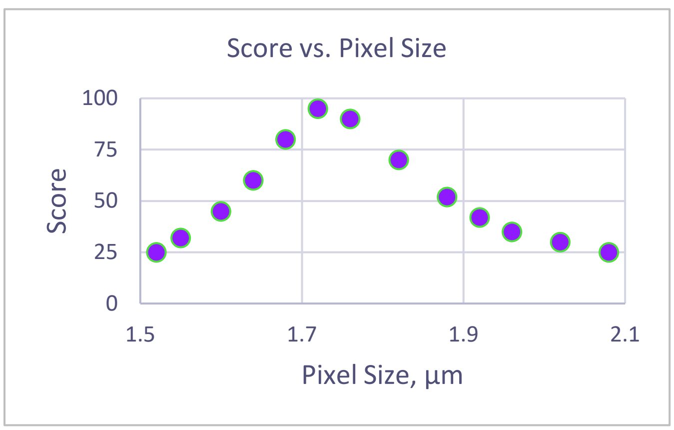


Figure 11. Edge pattern recognition score as a function of pixel size change (e.g.: scale difference between alignment model and image)

Image Size

The software will limit the minimum and maximum model size. A model that is on the larger range of size will typically perform better because there are more pixels available for matching.

Pattern Recognition Search Parameters

Several pattern recognition parameters are available to optimize pattern match and search. If the pattern is not found within the initial field of view, the system has the option to search for a pattern, which is called Groping. Multiple matching scores can be defined to help optimize the speed versus accuracy of pattern recognition.

Number of Groping Layers

The Number of Groping Layers determines the search area for the pattern recognition model. Each layer consists of a rectangular matrix of areas surrounding the deskew site. Each rectangular area is slightly smaller than the camera field of view to help ensure that the pattern is not lost between area boundaries.

None will only search for the pattern in the initial field of view. 1, 2, 3, and 4 Groping Layers will search 8, 24, 48, and 80 areas, respectively, around the initial field of view. This parameter should be set to account for sample placement error. It is recommended to set it to 1 Groping Layer to account for (1) typical manual loading error with precision locating pins for the P-7 and P-17, and (2) handler loading error for the P-170 and HRP-260.

Matching Scores

There are two scores that control when pattern recognition is complete: Minimum Score to Stop Groping and Minimum Match Score. The Minimum Score to Stop Groping allows the user to set a score threshold for successful pattern recognition, such that the system stops searching additional areas in the Groping Layer(s). The Minimum Match Score sets the lowest pattern recognition score that is successful after all areas have been searched in the Groping Layer(s).

Pattern recognition deskew will be used as a match score example. After the sample is loaded on the stage, the system moves to the first pattern recognition site to focus and adjust the light level (optional). A pattern recognition score will be generated. If the score is greater than or equal to the Minimum Score to Stop Groping, then pattern recognition is successful. If not, the stage is moved to the next area in the Groping Layer(s) and a new pattern recognition score is generated. This process is repeated until a score is equal to or greater than the Minimum Score to Stop Groping, or when all areas in the Groping Layer(s) have been searched. If Groping is complete without a successful match, the area with the maximum score is determined. If the score is greater than or equal to the Minimum Match Score, then pattern recognition is successful.

This process is then repeated at the second deskew site. If the pattern recognition score is less than the Minimum Match score, then pattern recognition has failed.

Pattern Recognition Matching Scores

The default scores in a new recipe may need to be adjusted to optimize performance for the sample and pattern recognition type that is being used. When setting the scores, there are two items that need to be considered: (1) the pattern recognition model must be unique within the Groping Layer(s) search area, and (2) a high pattern recognition score does not guarantee good pattern recognition performance. The difference in scores when the model is inside and outside the field is a better measure of pattern recognition performance.

A Best Known Method (BKM), "Pattern Recognition Model Optimization"[1], is available that describes how to optimize the scores a sequence recipe. Figure 12 shows the results of running this BKM on many different types of wafers using each pattern recognition matching method. The data shows that Geometric (GMF) performs the best, generally providing the highest difference in matching scores, but not for all samples.

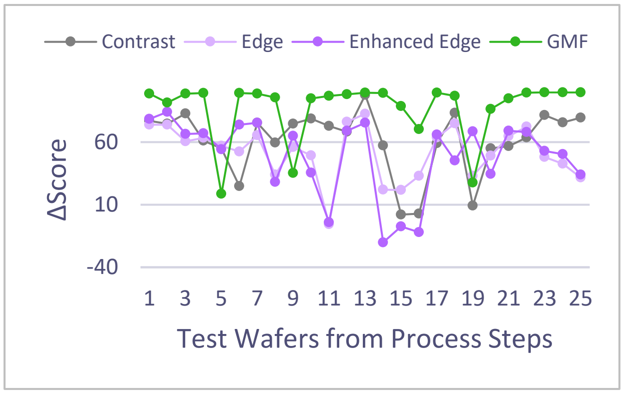


Figure 12. Pattern Recognition Model Optimization BKM results as a function of different wafer types and pattern recognition matching methods

Pattern Recognition Failure Options

If pattern recognition is not successful, the user can specify the next action. The most common action is to enable Operator Assist, where the operator can move the pattern into the center of the field of view to retry pattern recognition. The operator can also manually tell the software where the pattern is located to skip trying pattern recognition. Finally, the operator can skip the failed wafer or cancel the entire sequence recipe if pattern recognition fails. The software can also be set to automatically skip the wafer or cancel the entire sequence without operator intervention. It is also possible to continue measuring the scan site without successful pattern recognition, but this option is not typically used for a production environment.

Summary

Pattern recognition is a critical component to fully automate sample measurements. The sequence recipe on the Tencor P-7, P-17, P-170 and HRP-260 stylus profilers includes settings that enable customization of pattern recognition for different types of semiconductor processes such as plating, film deposition, etch, metal and oxide CMP, plus backside grind. This automation also ensures that each sample is measured with the same method, removing any operator error. For a high volume manufacturing environment, pattern recognition is critical to maximize tool utilization. For production and R&D, pattern recognition allows engineers and operators to perform other tasks while the profiler is measuring the samples. Pattern recognition, when combined with automated data analysis, SECS/GEM, and a handler, enables fully automated measurements.

References

[1] Best Known Method: Pattern Recognition Model Optimization. KLA Corporation, 2003.

KLA Corporation
One Technology Drive
Milpitas, CA 95035
Printed in the USA
Rev 9 2021-06-14