AIChE

Monitoring Roughness and Edge Shape on Semiconductors Through Multiresolution and Multivariate Image Analysis

Pierantonio Facco, Fabrizio Bezzo, and Massimiliano Barolo
Dipartimento di Principi e Impianti di Ingegneria Chimica, Università di Padova, via Marzolo 9, 35131 Padova (PD), Italy

Rajib Mukherjee and José A. Romagnoli
Gordon A. and Mary Cain Dept. of Chemical Engineering, Louisiana State University, Baton Rouge, LA 70803

DOI 10.1002/aic.11733
Published online March 27, 2009 in Wiley InterScience (www.interscience.wiley.com).

Photolithography is one of the most important processes in the production of integrated circuits. Usually, attentive inspections are required after this process, but are limited to the measurement of some physical parameters such as the critical dimension and the line edge roughness. In this paper, a novel multiresolution multivariate technique is presented to identify the abnormalities on the surface of a photolithographed device and the location of defects in a sensitive fashion by comparing it to a reference optimum, and generating fast, meaningful and reliable information. After analyzing the semiconductor surface image in different levels of resolutions via wavelet decomposition, the application of multivariate statistical monitoring tools allows the in-depth examination of the imprinted features of the product. A two level nested PCA model is used for surface roughness monitoring, while a new strategy based on "spatial moving window" PCA is proposed to analyze the shape of the patterned surface. The effectiveness of the proposed approach is tested in the case of semiconductor surface SEM images after the photolithography process. The approach is general and can be applied also to inspect a product through different types of images, different phases of the same production systems, or different processes. © 2009 American Institute of Chemical Engineers *AIChE J,* 55: 1147–1160, 2009
Keywords: process monitoring, semiconductor manufacturing, multivariate statistical methods, multiresolution methods

Introduction

Photolithography is a process that selectively removes parts from a thin film using light, so that a geometric pattern can be transferred (often from a mask) to a light sensitive chemical (the resist) deposited on a substrate. This process is used during the fabrication of integrated circuits (IC) as well as in many other micro-fabrication processes (e.g., micro-compressors production[1]). In particular, a microelectronics manufacturing process comprises an extensive sequence of complex semi-batch processes, among which photolithography is referred to as one of the most important.[2] In fact, photolithography: *i*) recurs up to 35 times for a given device; *ii*) defines the wafer critical dimension (CD) and the other

Correspondence concerning this article should be addressed to J.A. Romagnoli jose@lsu.edu.

© 2009 American Institute of Chemical Engineers

most influencing parameters; and *iii*) affects all the successive processing phases (e.g., the doping) and the interconnection between different segments of the device. From an economical point of view, the lithography is responsible for about 60% of the processing time and 35–40% of the total cost of the IC fabrication.[2] As a consequence, it is quite clear that monitoring the product quality during photolithography through a fast, sensitive, and reliable system is highly advocated.

Although considerable effort has been dedicated to define technologies and procedures to meet the requirements on the product quality,[3,4] automatic process control has not yet been implemented on a large scale in semiconductor manufacturing, and the industrial practice is often carried out empirically with relatively little understanding of the underlying physics and chemistry.[5] Statistical process control techniques, too, are sometimes adopted[5–7] in order to monitor the variability of the process, to detect the abnormal conditions, and to identify the cause for a perceived anomaly.

Currently, the most advanced monitoring strategies exploit hardware and software devices for both signal filtering and image processing.[8,9] For instance, the use of scanning electron microscopy (SEM) images is common for the measurement of the physical parameters of a device[10] such as the CD.[11] However, the typical inspecting tools focus on inline optical metrology systems measuring the CD of the pattern and its variability; only the most sophisticated instruments also determine the edge height and the side-wall angle[12] (SWA). Several important quality features like the line edge roughness (LER), the edge surface roughness, the actual shape of an edge (and its variability) are still rather resilient to effective, fast and low-cost monitoring technologies. Only recently some researchers[4,13] have suggested procedures to start tackling the above issues.

Thus, the demand of satisfying the multiple requirements of the wafer fabrication and the dynamics of a quickly changeable microelectronics market call for new and more powerful instrumentation. The accuracy of the manufacturing quality monitoring can be greatly improved if fast and more meaningful information were retrieved in a reliable fashion.

One promising approach comes from suitably exploiting the amount of information contained in SEM images. For example, they may allow for a simultaneous and more probing examination of the complex device cross-section, of the pattern shape, and of the surface roughness. Indeed, the photolithographed product analysis can be improved considering that the images are multivariable signals in nature, and that multivariate statistical techniques are some of the most suitable and powerful methods to inspect a multivariate system.[14]

In this work a novel and general methodology is presented, where a multi-resolution filtering technique and multivariate image analysis are combined to inspect the surface of a semiconductor after photolithography. The main components of the proposed quality monitoring strategy are:

- sensitive filtering pre-treatment, which de-noises the image signal and removes the image artifacts without affecting the featured parts and their peculiar characteristics (i.e., the surface roughness);
- tailored multivariate statistical models, based on a principal component analysis[15] (PCA) approach, which extract the information content embedded in such composite signals; namely, quality information concerning LER, surface roughness and edge shape will be obtained.

The article is structured as follows. After a brief introduction to the photolithography process, the proposed image analysis technique will be described. A section will then detail the approach for multi-resolution denoising and why this step is needed. The subsequent section will explain the multivariate monitoring methods developed in this work and comment on their applicability and advantages. A case study of an after-photolithography inspection will be used to illustrate the effectiveness of the methodology. Some final remarks will conclude the paper.

Photolithography Process and Inspection Tools

Integrated circuits are recognized as some of the most complex manufactured products and some of the most versatile devices. As a basis, IC fabrication is obtained through a complex infrastructure of materials supply, waste treatment, logistics, and automation to support the entire process. Specifically, the semiconductor production technology develops through an extensive series of photographical, mechanical, and chemical steps in the cleanest environment, achieved by ultra-precision engineered equipments.[16] Typical processing loops, which may recur several times, comprise some or all of the following phases (Figure 1a): oxidation; photoresist application; exposure to light; development of the resist; etching; and photoresist removal.

In detail, the abovementioned procedure can be outlined as follows. First of all, the wafer, a thin crystalline slice from a semiconductor ingot (e.g.: silicon), is heated to drive off the moisture from the surface, and cleaned. Later, the wafer is maintained into a high temperature environment, until an oxide layer grows on the substrate. After the addition of adhesion promoters, a thin (and as uniform as possible) layer of photoresist is applied by spin-coating, in a high-speed centrifugal whirling process. Essentially, the photoresist is a polymer mixed with light sensitive compounds. Through light exposure, the desired pattern (often determined by a mask) can be impressed on the surface illuminating certain portions of the resist selectively.

If the light weakens the polymer to the so called "positive" photoresist, it will becomes a less chemically stable aggregate and will be more easily removed during the following stages. Conversely, a "negative" photoresist is strengthened by light and becomes resistant to solvents. The chemical change triggered by the light during the photolithography step allows the resist to be removed by a solution called developer. The resulting shape of the device surface should be the one shown in Figure 1a, alternating zones in which photoresist is present, the so called *edges*, and zones (which in this paper will be indicated as *valleys*) in which the oxidized substrate is no more protected and is completely free from the resist (Figure 1b).

After the development, a hard baking is performed to give a stronger structure to the residual resist and only then, during the etching, the part of the surface that is not protected by the resist is engraved. A chemical agent (a liquid or a plasma) removes the oxide layer to prepare the surface for the following phases. Finally, the remaining resist is

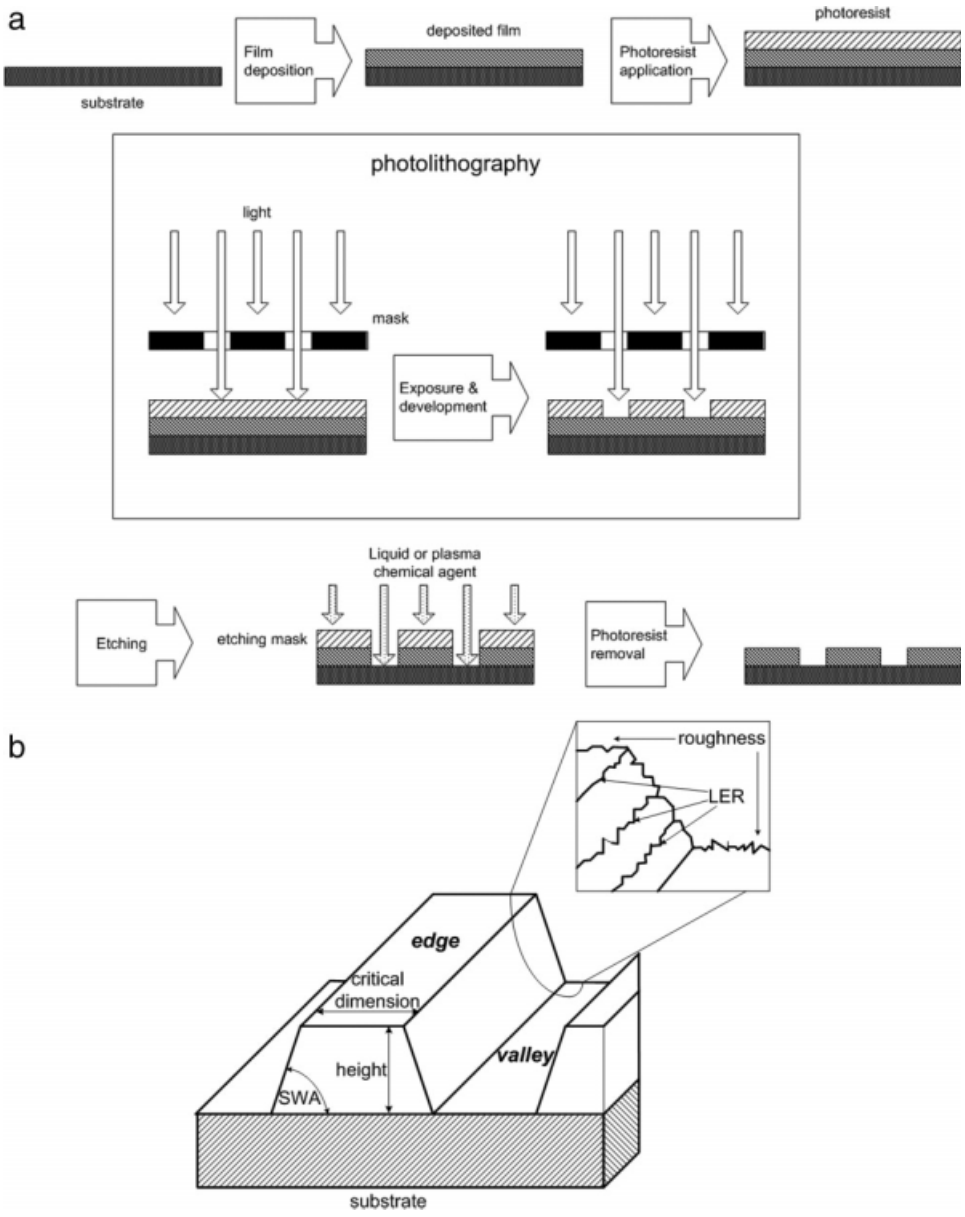

Figure 1. (a) Simplified processing sequence for the semiconductor manufacturing and (b) simplified graphical representation of the most important quality parameters of an edge.

removed and the surface is ready for the diffusion of dopants on the part of the surface where the oxide barrier is not present. The doping induces the formation of ions that create regions with different electrical properties.

During each loop, it is crucial to meet stringent requirements in terms of quality uniformity and consistency. At each stage, several inspections and measurements are performed, but they monitor only few samples of different lots, and some pieces of the processing equipment. Since photolithography is performed several times on the same device and, even after it is completed, a defective device can still be reprocessed, it is common industrial practice to perform quality inspections after the photolithography step (the so-called after-development inspections). Usually, a CD-SEM is adopted to measure some significant features of the semiconductor surface. The SEM images are used mainly for metrology purposes. In other words, the common inspection tools measure the CD or, in the most advanced instrumentation, the edge height, the SWA, or the LER (Figure 1b). Recently, the possibility of measuring and reducing the edge wall roughness using deep UV light scatterometry has been discussed.[4] However, in order to detect, distinguish and classify critical features of the manufactured device, more sensitive and reliable tools are required by the new generation of products.[3] For instance, there is a number of defects affecting the final product quality and performance (Figure 2) that cannot be identified in terms of CD measurements, but to which it could be possible to remedy if timely detected.

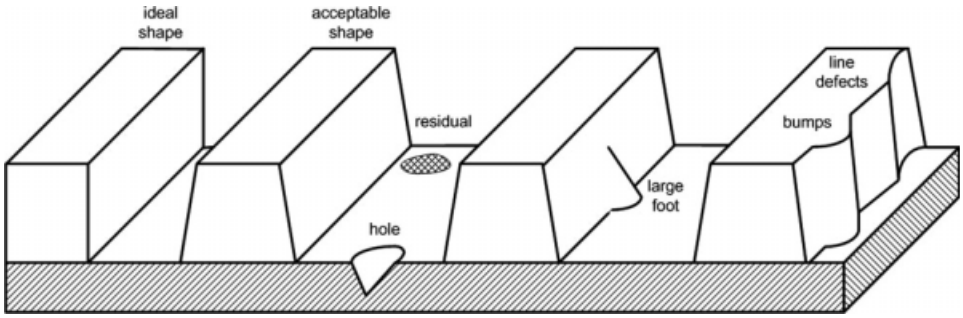

Figure 2. Ideal shape of an edge and possible defects on the edge shape and surface.

Thus, an automatic monitoring system for the frequent and accurate quality monitoring of a photolithographed surface would be a highly attractive perspective to increase the yield and the consistency of the fabrication program.

Image Analysis through Multiresolution and Multivariate Statistical Techniques

Images are bi-dimensional (2-D) maps summarizing the characteristics of a three-dimensional (3-D) scene. In this work, industrial SEM images representing the product surface after photolithography with positive photoresist will be used as a case study. These images are grey scale functions of light intensities and can be used to extract meaningful information about the quality of the product, its regularity and conformity to the requirements, and the types and location of defects. The scale of the images is such that 1 pixel corresponds to about 9.8 nm.

In general, an image is a collection of well identified characteristics. As such, it is intrinsically a multivariate system, being a wide collection of pixels, where each pixel is highly correlated to its neighbors. In addition, in a surface image the apparent variability comprises both the actual roughness of the surface (which is an actual product feature) and the signal disturbance (which corresponds to noise to be removed). Thus, a number of tasks need to be taken into account by a monitoring system in order to ''use'' an image for quality control. Figure 3 illustrates the general architecture for a monitoring system.

First of all, a reference model is defined by selecting a suitable reference image. As the quality problem involves different scales of resolution, a multiresolution approach is needed to filter the image and denoise the signal. Subsequently, multivariate statistical techniques are used to exploit the information content of the filtered image, to formulate the monitoring model, and to build the monitoring charts for the product quality inspection. Three quality features are described and monitored through the monitoring system: the LER, the surface roughness, and the shape of an edge trans-section profile.

In the following subsections, the main properties and the mathematical foundations of the multiresolution-multivariate monitoring system will be discussed.

Image multiresolution denoising

An image is always affected by disturbances, e.g., the random fluctuation of the pixel light intensity. In general, multivariate statistical techniques can discard the nonsystematic part of a signal, distinguishing the meaningful variability from the random one. Unfortunately, in this case, the noise is somehow blurred with the roughness of the surface (which is a structural part of the device and defines the quality properties one is interested in). Therefore, a pretreatment is needed on the image to remove the noise without discarding the structure of the surface roughness.

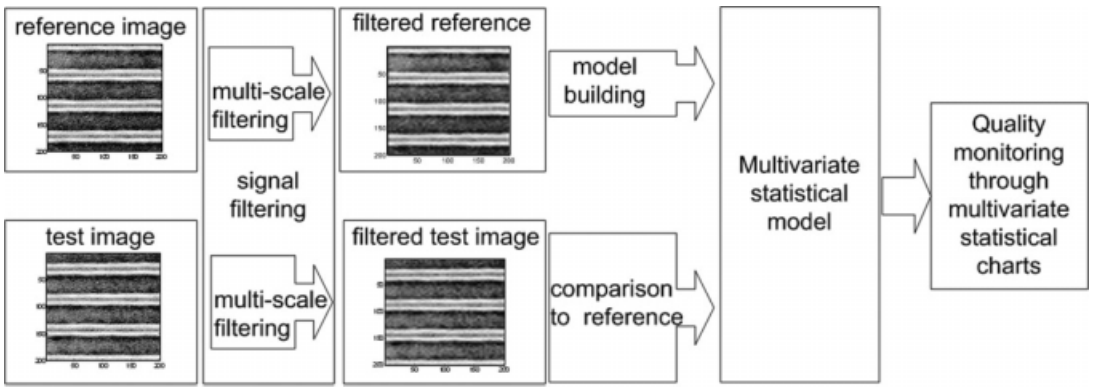

Figure 3. Sketch of the semiconductor monitoring system through (wavelet) image filtering and multivariate statistical techniques.

The problem of the dual nature of the noise is faced following a multi-resolution approach, which examines the different scales of the image through wavelet decomposition (details on wavelets can be found in Refs. 17 and 18). Specifically, a scale-dependent smoothing of the image is performed by subtracting the unwanted part of the noise. In fact, the smoother version i_M of the image i in the domain of the pixel space $\mathbf{s} \in \Re^2$ is the approximation at the scale M of the original image:

$$i_M(\mathbf{s}) = \sum_{n=-\infty}^{+\infty} S_{M,n}\phi_{M,n}(\mathbf{s}) + \sum_{m=-\infty}^{M}\sum_{n=-\infty}^{+\infty} T_{m,n}^{\text{denoised}}\psi_{m,n}(\mathbf{s}), \quad (1)$$

where the first term on the right-hand side is the summation of the products between the approximation coefficients $S_{M,n}$ at the M^{th} scale and the selected father wavelet ϕ (lower frequencies of the signal), and the second term is the summation of the products between the detail coefficients $T_{m,n}^{\text{denoised}}$ and the mother wavelet ψ (higher frequencies of the signal). Some of the higher frequencies (scales m over a prescribed limit M_1) are redundant and, as a consequence, can be removed:

$$T_{m,n}^{\text{denoised}} = \begin{cases} 0 & m \geq M_1 \\ T_{m,n} & m < M_1 \end{cases}. \quad (2)$$

Only the lower resolutions are retained, because of their relevance to the purpose of the roughness monitoring. So, after being convolved through the wavelet decomposition in different scales of resolution, the image is reconstructed, by merging together all the significant scales.

Different types of wavelets were tested for the denoising of the photolithographed surface. The Daubechies wavelet with 8 scaling coefficients was eventually selected. The use of the Daubechies wavelet is suggested in several studies,[13,19] especially for segmentation and texture analysis problems. Indeed, the Daubechies 8 wavelet seems to respond very well to the requirements discussed in Ref. 20 as it introduces very limited phase distortion, maintains a faithful localization on the spatial domain, and decorrelates the signal in a sensitive manner for both the image smooth features and discontinuities.

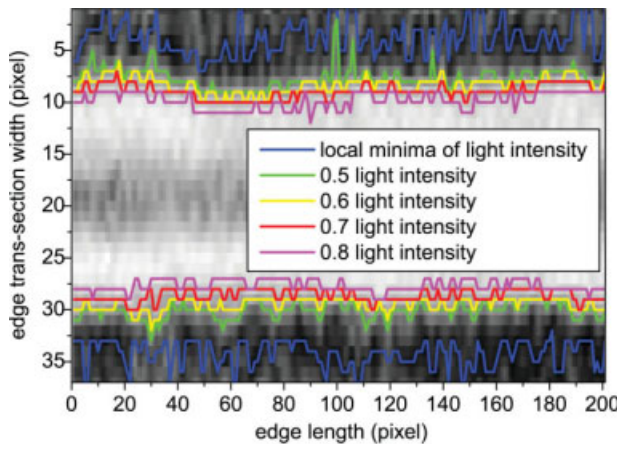

Figure 4. Magnified section of an edge image: detection of topological levels at different light intensities for the identification of the side-wall roughness on a reconstructed Daubechies-8 1st level approximation.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

Table 1. Correlation Coefficients Between Different Topological Lines in the Original Image and in the 1st and 2nd Level Daubechies 8 Approximation Reconstructions

	Threshold	0.5	0.6	0.7	0.8
Original image	0.5	1	0.5974	0.3881	0.3316
	0.6	0.5974	1	0.5030	0.4573
	0.7	0.3881	0.5030	1	0.6179
	0.8	0.3316	0.4573	0.6179	1
1st level approximation	0.5	1	0.7506	0.7676	0.5167
	0.6	0.7506	1	0.7164	0.6321
	0.7	0.7676	0.7164	1	0.6761
	0.8	0.5167	0.6321	0.6761	1
2nd level approximation	0.5	1	0.8086	0.7971	0.7784
	0.6	0.8086	1	0.8337	0.7787
	0.7	0.7971	0.8337	1	0.8889
	0.8	0.7784	0.7787	0.8889	1

Once a reference wavelet has been identified, the issue is to find the best smoothing scale for roughness monitoring. We found that useful indications can be obtained by evaluating the correlation coefficients between the side-wall roughness lines at different light intensity levels. A procedure inspired by the work of Patsis[21] was followed to detect the side-wall roughness. The locus of minimum light intensity in the valleys (represented by the blue lines in Figure 4) is first recognized, and then one moves upwards along any of the two edge walls (*upper* wall and *lower* wall*), detecting different light intensity levels at preset thresholds. In this way, the "topological lines" at light intensity thresholds of 0.5, 0.6, 0.7, and 0.8 are identified (respectively the green, yellow, red, and violet lines in Figure 4). Stated differently, these lines represent the pixel location along an edge side wall where the light intensity assumes a certain value (iso-intensity lines). Therefore, according to this procedure, a spatial location along the edge wall is identified through a certain light intensity level.

The correlation coefficients between the different topological lines are computed in the cases of the original image and of the reconstructions at the 1st and at the 2nd approximation levels. As can be seen in Table 1, the correlation coefficients between the topological lines of the original image are always low (even for neighboring lines). This means that the noise corrupts the image determining an artificial "uncorrelation" among pixels. However, note that the reconstruction at the 1st approximation level exhibits a substantially different situation: higher correlation is shown between neighboring lines and a significantly lower correlation exists between distant lines. This occurrence is related to a certain level of roughness, which determines a random shaping of the edge wall (hence, of the topological lines). If the filtering level is further increased as in the case of the reconstruction image

*The definition of upper and lower walls simply refer to the order they appear in our images, i.e. for one edge the upper and lower wall are respectively the first and second wall when moving in a top-down direction. In principles, they do not have any systematic (physical) difference; however, from a statistical point of view, we observed that they "belong" to different categories (perhaps because the light hit them from a slightly different angle) and therefore we decided to identify them.

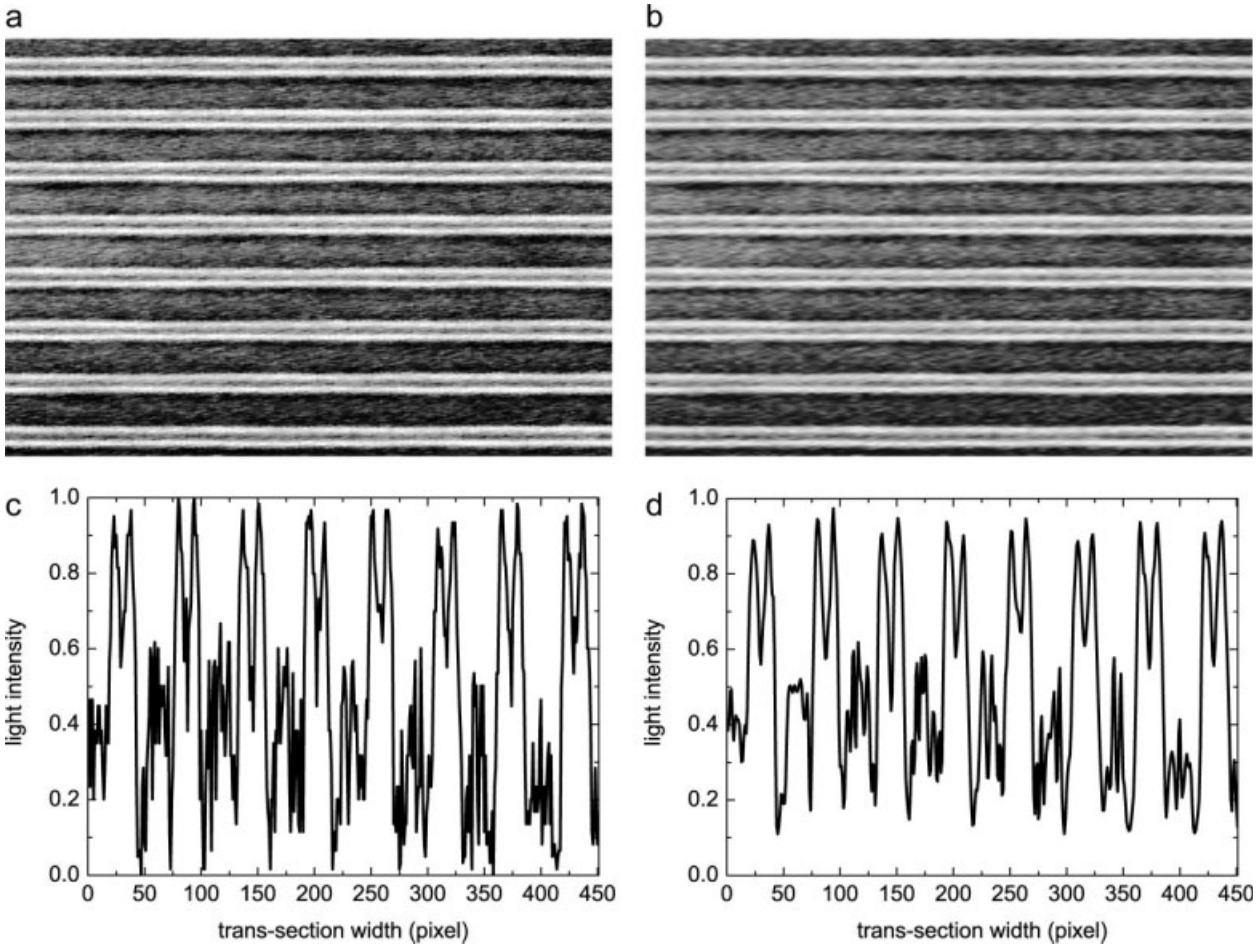

Figure 5. (a) Original image and (b) filtered image, with an example of trans-section profiles on the same pixel column for (c) the original image and (d) the filtered one.

at the 2nd approximation level, the random shaping of the lines is excessively reduced, i.e., the ability to capture the roughness is lost. In fact, the correlation coefficient between far thresholds (0.5 and 0.8) is almost the same as the one between neighboring thresholds. This means that an excessive ''smoothing'' of the signal has caused both the noise and the structural roughness to be removed from the original image. Thus, a first level decomposition is chosen.

After normalization (i.e., forcing the spectrum of the black-and-white light intensities between the values 0 and 1) and wavelet filtering, the resulting smoothing on the original image is shown in Figures 5a,b. The trans-section profiles of the original image along a given pixel column (Figure 5c) show a confused trend (the noise is almost indistinguishable from the underlying pattern). Once the image is filtered and polished from the high frequency components, it gets a clearer and more noticeable pattern (Figure 5d), to which multivariate statistical techniques can be applied.

Multivariate statistical surface monitoring methods

Multivariate monitoring instruments are needed for analyzing and interpreting the multivariate nature of images. Whenever there is the technical impossibility to perform a measurement, or an analysis entails the simultaneous accessibility of multiple characteristics, a multivariate statistical monitoring system may prove significantly more powerful and effective than common metrology tools.

Different multivariate statistical schemes were adopted through modified PCA approaches. An extended treatment on PCA methods can be found in textbooks.[14,22] The basis for a multivariate monitoring technique is its capability of summarizing a plurality of quality clues embodied in an image into a limited number of statistical indices. Usually these are an indicator of the mean trend (the Hotelling T^2 statistic, as a replacement of the scores $\mathbf{t}$), and an indicator of the model representation suitability (the Q statistic).

As stated before, the objective is to develop an image-based strategy for the monitoring of LER, surface roughness and edge shape.

LER monitoring

Line edge roughness is one of the typical reference parameters in the after-development inspection of a photolithographed device.[4] In fact, it plainly affects both the subsequent production phases and the performance of the final product.

To monitor the LER of a single edge, information must be provided on how the light intensity is distributed in the upper and lower walls of the edge. As was mentioned in the

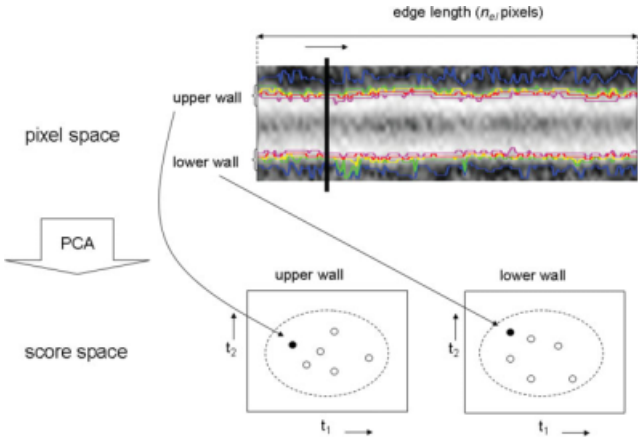

Figure 6. Inspection of the upper and lower walls of an edge by the LER monitoring model.

Two latent variables were used in the PCA model. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

previous section, the topological lines on $n_{\text{levels}} = 4$ light intensity levels were identified. Two PCA models[23] were developed, one for the upper wall and one for the lower wall. To build each model, a reference $[n_{\text{el}} \times n_{\text{levels}}]$ data matrix was used, where n_{el} is the edge length (in pixels). Each column of this matrix represents the pixel locations, along an edge trans-section (identified by the relevant row number), where the light intensity assumes the value of 0.5, 0.6, 0.7, or 0.8. Otherwise stated, each column represents the topological line where light intensity assumes a specified value.

For each trans-section of an edge, the PCA model combines the n_{levels} variables into a measure of the mean and variance of the LER, and produces one point in the scores space representing the edge quality on that trans-section of the edge. When a new edge is inspected, a trans-section is considered (thick black line in the upper part of Figure 6), and the relevant topological levels data are projected onto the reference model. Differences in the mean of topological lines at a given light intensity (e.g., due to the presence of side-wall bumps) or in their variance (e.g., due to the presence of large feet or spikes) are highlighted by the T^2 (or scores) and the Q monitoring charts (respectively). By proceeding this way for each trans-section, the whole edge length can be scanned and inspected. This strategy allows to detect and localize the imperfections on the edge side walls, which can then be related to malfunctions or drifts in the production machinery. The confidence levels for the T^2 and Q thresholds can be selected in such a way as to guarantee a good sensitivity to faults, while at the same time keeping the number of false alarms small.

Surface roughness monitoring

Semiconductor surface roughness is also a very important parameter, as it can deliver substantial information concerning the accuracy of the erosion during photolithography, the presence of resist residuals on the substrate, etc. The challenge in monitoring is that the patterned surface of an after-photolithography photoresist shows uneven characteristics in different positions. In fact, the zones at lower light intensity (valleys) have remarkably larger roughness than the zones at higher light intensities (edges), due to irregular erosion from the light. Therefore, not only the edge surface should be inspected, but also the valley surface must be analyzed. This also implies that one must be able to automatically distinguish between edges and valleys along the whole surface of the device.

A single model on the entire semiconductor surface cannot simultaneously capture the uneven variance structure on different locations of the surface. It has been suggested[13] to segment the surface in classes through a k-means clustering on the light intensity. We propose here to use an unsupervised PCA discriminant analysis to distinguish the edges from the valleys. As a result, for each surface configuration (edge or valley) a specific monitoring model is built.

A nested PCA monitoring system was designed to this purpose, which is based on a two-stage procedure: the outer level carries out the unsupervised discriminant analysis; the inner level is the actual monitoring model, where two PCA submodels are enclosed, one for the edges and one for the valleys. Thus, in the sequential surface scanning procedure, the outer level provides an automatic switching from one submodel to the other, and the correct submodel is interrogated in the inner level for the monitoring step.

Each submodel works on either an $[n^{\text{E}}_{\text{iw}} \times n_{\text{el}}]$ or an $[n^{\text{V}}_{\text{iw}} \times n_{\text{el}}]$ matrix, where n^{E}_{iw} and n^{V}_{iw} are the reference edge and valley image widths (in pixels), respectively. The reference images for edges and valleys are built by considering a statistically sound number of on-quality edges and valleys (in fact, it is sufficient that a statistically sound number n^{E}_{iw} and n^{V}_{iw} of on-quality edge and valley image rows is collected and assembled in two reference images). The j-th column of each matrix contains the values of light intensities on the image trans-section located at pixel $0 \le j \le n_{\text{el}}$ along the image (i.e., edge) length. The PCA submodels reduce the dimension of the problem (n_{el} variables) into a point on a 2-D scores space that characterizes most of the information content and the variability of the original space. As

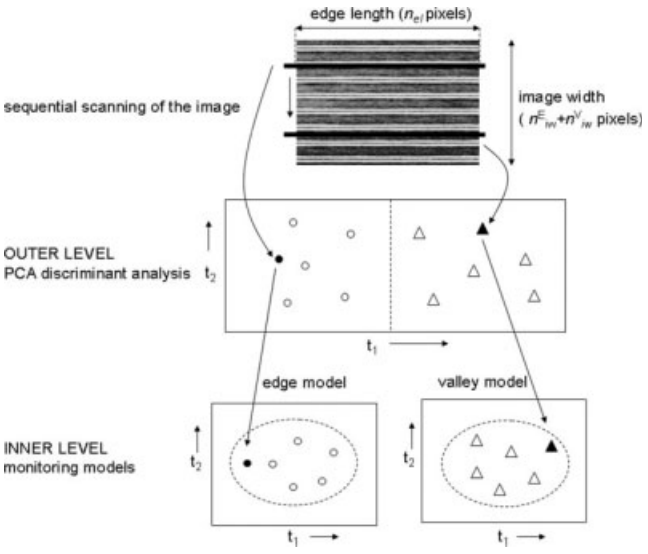

Figure 7. Edge and valley surface roughness monitoring using the nested PCA model.

Two latent variables were used.

illustrated in Figure 7, a new picture is scanned row by row; any new row is identified as belonging either to an edge or to a valley, and then transformed by using the ad-hoc PCA model. The value of the first score t_1 was found to be a useful indicator to classify a row. A threshold value for t_1 was determined by analysis of the reference images.

Each submodel allows to monitor the surface roughness by analyzing the T^2 and the Q indices. Excessively high values for the Hotelling's T^2 statistic alarm on an abnormal mean light intensity along the whole length of an edge or valley (e.g., due to uneven distribution of the photoresist coating). Excessively high values for the Q statistic indicate an abnormal variability of the light intensity (e.g., due to the presence of holes or of photoresist residuals). Note that any abnormality can be precisely located by analyzing the contribution provided by each of the n_{el} pixels to the altered value of the relevant statistic. Therefore, the information that becomes available through this surface monitoring system complements and extends the information provided by the LER monitoring system.

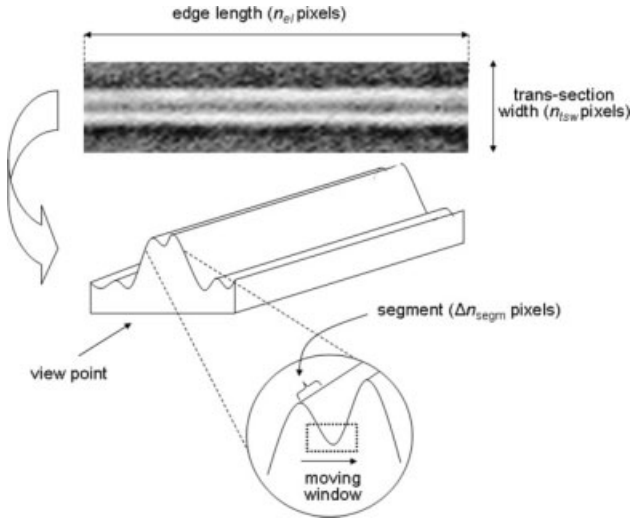

Figure 8. Schematic of the main concepts underneath the monitoring method for the shape analysis of one edge.

Edge shape monitoring

A more wholesome and meaningful monitoring approach would attempt to examine the overall edge shape and to compare it to a required standard. A methodology is suggested here to achieve this goal. The proposed approach performs the inspection of a single edge shape by scanning the edge image in the sense of the pixel columns. The objective is to compare different trans-section profiles of an edge (as characterized in terms of light intensity) with respect to a reference profile, while proceeding along the edge length direction.

There are two main difficulties in monitoring the edge shape. First of all, the data are clearly nonlinear and non normally distributed in the direction of the image column trans-section, whereas the calibration of a PCA model requires linear and normally distributed historical reference information. Secondly, it is difficult to retain the spatial information when using a PCA model as, ideally, every pixel should be considered for its position and for the spatial relation with its neighbors.[24]

The strategy proposed to overcome these issues is to consider each pixel together with its nearest neighbors located in the same trans-section profile within a predefined spatial moving window (Figure 8), in which the correlation between the neighbors (i.e., the local edge shape) is maintained and the nonlinearity of the profile is negligible. Thus, a moving window of appropriate size (Δn_{mw}) is defined in the space of the pixels along the edge profile trans-section (whose overall width is n_{tsw} pixels). The moving window captures the correlation between pixels located in a limited region, within which it is reasonable to suppose a strong correlation between neighbors. For this reason, $\Delta n_{mw} = 5$ pixels was chosen; at larger distances the correlation between upper and lower pixels within the same window vanishes.

In order to smooth nano-variations that may perturb the comparison between the local shapes of different edge profiles, a segment is taken along the edge length direction (a segment dimension of $\Delta n_{segm} = 5$ pixels was chosen for this study, which corresponds to $\sim$49 nm; at larger distances the correlation between right and left pixels was not significantly high). This segment is assumed to be the smallest width for which a profile shape is analyzed; this also allows for a significant reduction in the computational burden. Therefore, each segment represents the image of Δn_{segm} consecutive edge profiles. The shape of an edge is then analyzed by monitoring the segments that edge is made of. Note that a segment can be represented by a $[n_{tsw} \times \Delta n_{segm}]$ matrix, whose entries are the light intensities on each pixel belonging to the segment.

To define the optimal reference for a segment, a number (equal to N_{image}) of segments that conform to the required quality standards are collected from different edges and different images. The whole set of reference data is arranged in a 3-D matrix (Figure 9), whose dimension is $[N_{image} \times \Delta n_{segm} \times \Delta n_{tsw}]$. Therefore, according to this arrangement, the N_{image} images are stored as horizontal slices and piled up one another.

The spatial moving window multi-way PCA model is then defined as follows. The spatial window moves pixel by pixel along the matrix third dimension. For any position of the window, a subset of the 3-D matrix is defined (Figure 9a). This subset is unfolded ''image-wise'', by cutting the submatrix into Δn_{mw} vertical slices along the trans-section width dimension, and putting the slices side by side according to a multi-way procedure[25] (Figure 9b). This results in a $[N_{image} \times (\Delta n_{mw} \cdot \Delta n_{segm})]$ 2-D matrix that can be processed through PCA. A column of this matrix represents how the light intensity on a given position in segment and on a given position along the edge profile varies between different images. On the whole, $(n_{tsw} - \Delta n_{mw} + 1)$ 2-D reference matrices are obtained, this number being equal to the number of positions that the spatial window can take along the edge trans-section width. For each of these matrices, threshold values on the T^2 and Q statistics can be determined that allow to monitor the shape of the edge within the corresponding spatial window.

For any segment on a test edge whose shape needs to be monitored, the T^2 statistic (or the scores) on a given edge window summarizes the mean edge shape in that window.

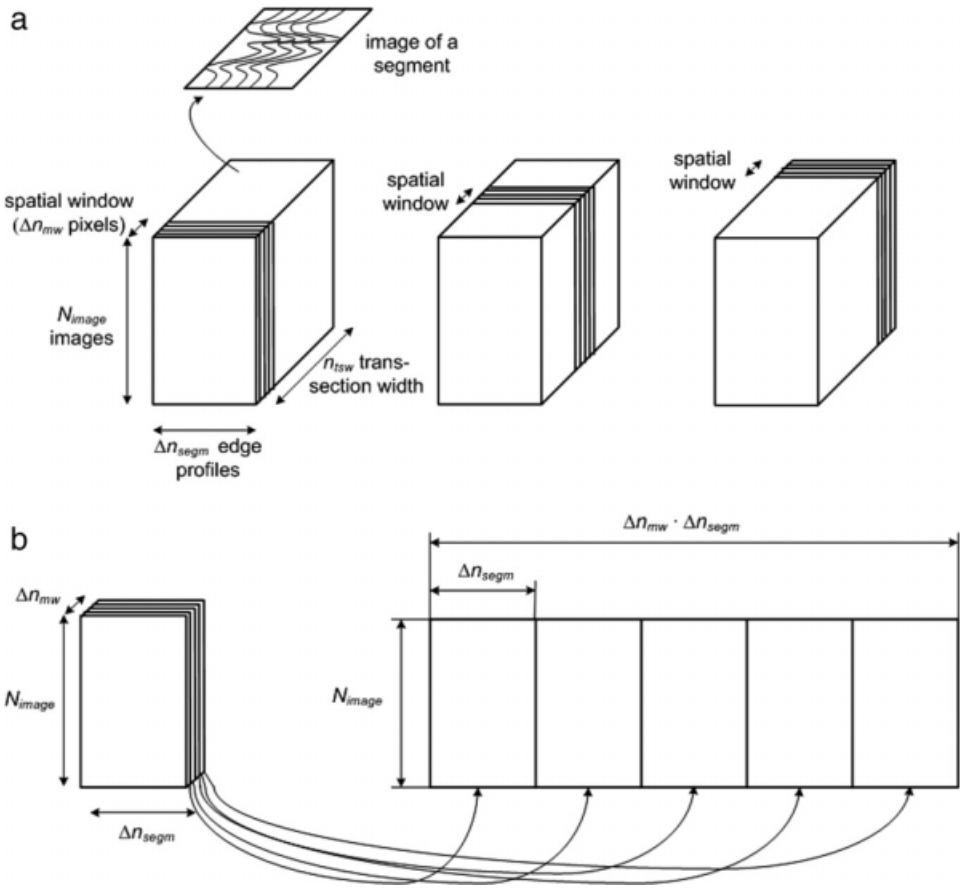

Figure 9. (a) Reference data arrangement on a 3-D matrix with a spatial moving window. (b) Image-wise unfolding of the 3-D matrix on the moving window.

Therefore, large T^2 values indicate that the local edge shape is altered with respect to the average. Large Q values indicate changes in the correlation structure between the profiles of a segment. This signals a local variation on the edge roughness, if the T^2 statistic is found to be within its limit.

Case Study: Monitoring Results

The compliance of the surface product quality to the quality requirements after a photolithography process is evaluated in terms of LER, of surface roughness and of edge shape by applying the techniques described in the previous section. Hence, a multiple monitoring system is developed to perform the monitoring of all the abovementioned features through multi-scale and multivariate image analysis. The monitoring results are presented in the following.

LER monitoring system

The LER monitoring strategy goes through the following steps: (i) an edge is selected on the de-noised test image; (ii) the four light intensity levels are identified on the edge; and (iii) the edge is monitored through the scores plot (and/or the T^2 plot) and the Q plot.

The scores plot surveys the conformity of the mean side-wall roughness on the edge. Any point in this plot is designated with a number, which represents the column position (along the edge length) to which that point refers to. The location of the point on the scores plot represents how the topological lines are distributed along the edge side-wall at that column position. Nonconformities are shown as points located outside the confidence ellipse.

The Q plot points to irregularities in terms of excessive variance on the edge side-wall roughness, and to changes in the correlation between the topological lines (i.e., in the "parallelism" of topological lines). The column location over the edge length is represented on the x-axis of this plot. Nonconformities are shown as points located above the confidence threshold.

In Figure 10, the scores plot (Figure 10a) and the Q plot (Figure 10b) for the upper side wall of an edge are shown. The Q plot shows four outliers within the first 30 pixel columns. An off-line visual inspection of the upper side wall (Figure 10c) indeed confirms that a bump is localized around column 30 (note that the scores plot, too, provides a "mild" alarm for column 30). Furthermore, Figure 10c shows that several side-wall irregularities are present in the first 30 columns, i.e., a nonparallelism of the topological lines exists in this section of the edge length. This confirms the indications provided by the Q plot. Note, however, that inspecting the side-wall image is much more time consuming and does not provide a precise and unambiguous indication on the pixel

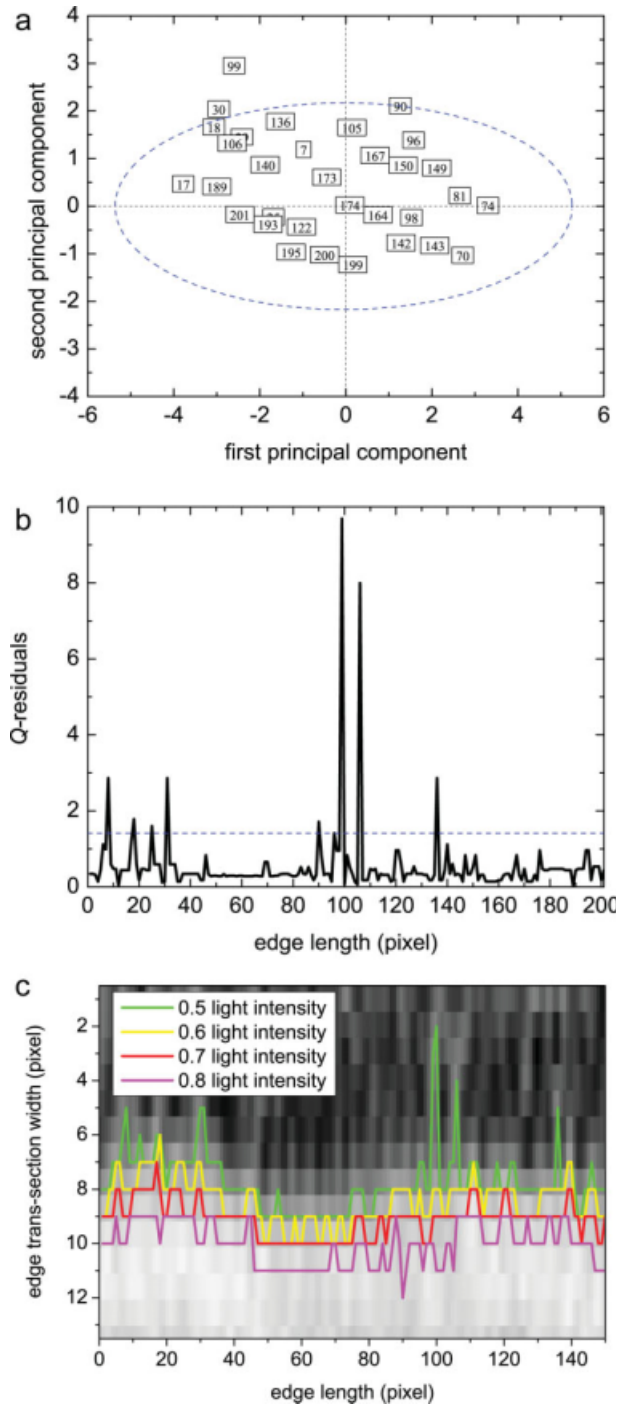

Figure 10. Line edge roughness monitoring: analysis of the upper side wall for one selected edge.

(a) Scores plot (the numbers within the squares represent the column position along the edge length); (b) Q plot; and (c) magnified section of the edge upper wall image. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

column where a nonconformity is present. Conversely, the Q chart is very quick to analyze, the response is localized and unambiguously points to a column where a defect is deemed to be present. The scores plot is somewhat less responsive than the Q plot, but nevertheless the information provided by the two monitoring diagrams can complement each other.

Strong outliers are also identified by the Q plot on pixel columns 99 and 106. These indicate a very large variability on the topological lines, or a nonparallelism between the lines. The irregularity on column 99 is also clearly detected by the scores plot. Visual inspection of the upper side wall image (Figure 10c) confirms that two large feet are present around pixel column 100.

Surface roughness monitoring system

The surface roughness monitoring strategy goes through the following steps: *i*) a row is selected on the de-noised test image following a sequential scanning from the top to the bottom of the image; *ii*) the row is categorized as belonging either to an edge or to a valley; *iii*) the edge/valley is inspected through the relevant scores plot (and/or the T^2 plot) and the Q residuals plot, which highlight potential anomalies on the surface; and *iv*) an analysis of the contribution of each pixel to the T^2 statistic and to the **e** residuals, which constitute the Q-residuals,[26,27] allows to precisely localize the imperfections.

Figure 11 shows the **e** contributions to the residuals of all the pixels along the length of a row within the same edge considered in Figure 10. If the **e** contribution of a pixel exceeds one of the confidence limits,[28] this is an indication that a surface imperfection is localized on that pixel. Following this rationale, several imperfections are detected by the contributions plot of Figure 11 around pixel columns 10 to 40, around column 100, and around column 135. Note that irregularities in the same locations were highlighted in the LER monitoring strategy, too.

The agreement between the results of LER monitoring and of surface roughness monitoring is a proof of the reliability of both approaches, which can be used simultaneously in a robust way. In fact, the topological lines at different light intensity thresholds (obtained by cutting the image

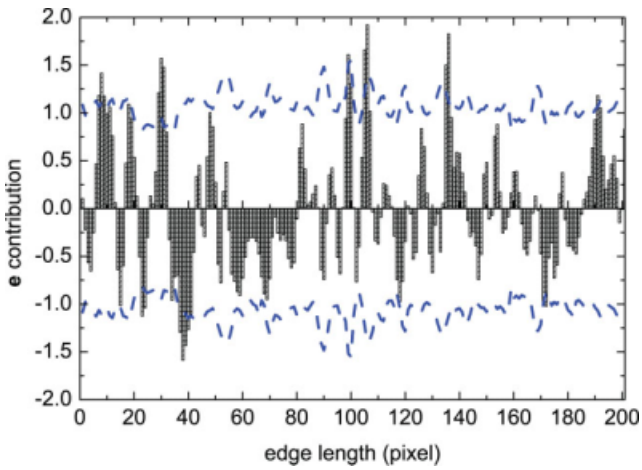

Figure 11. Surface roughness monitoring: typical trend of the contributions to the residuals for the localization of defects on an image row.

The dashed lines represent the confidence limits. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

Table 2. Correlation Coefficients Between the Topological Lines at Different Light Intensity Thresholds and the Light Intensity Along the Trans-Section of an Edge

	0.5 threshold	0.6 threshold	0.7 threshold	0.8 threshold
Row 5	0.432	0.245	0.138	0.098
Row 6	0.731	0.769	0.579	0.372
Row 7	0.700	0.866	0.728	0.516
Row 8	0.663	0.887	0.863	0.735
Row 9	0.525	0.752	0.841	0.829
Row 10	0.466	0.631	0.744	0.795
Row 11	0.216	0.205	0.285	0.381

with a plane parallel to the image itself) are highly correlated to the light intensities along the trans-sections of the edge surface (obtained by cutting the image with a plane orthogonal to the image). As an example, Table 2 shows the correlation coefficients between the upper side-wall topological levels of an edge and the trans-section profiles along the edge length from row 5 up to row 11. The higher correlation coefficients emphasize the similarity between the trajectories of the LER lines and the profile of the light intensity on a row of pixels along the edge length. This confirms the possibility to observe accurately the LER through the surface roughness in terms of light intensity.

Note that the surface roughness analysis can be used to spot also other defects (which usually cannot be identified in a practical way). For instance, Figure 12a shows the contributions plot for a different pixel row along the image (the row was categorized as a valley). It can be seen that around pixel columns 25, 50, 100, 160, and 195 the contributions to the residuals exceed the confidence limits. This means that some defects are present in terms of excessive variance of the light intensity (which could be associated to the presence of holes or of photoresist residuals). However, only a few of these defects would unambiguously be highlighted if an off-line inspection of the light intensity profile were carried out on the same row (Figure 12b).

Edge shape monitoring system

The edge shape monitoring strategy goes through the following steps: *i*) a segment is selected in the de-noised image of an edge; *ii*) the trans-section profiles in this segment are aligned to the reference one; *iii*) the segment is sequentially scanned by the spatial moving window along the whole trans-section width; and *iv*) the shape conformity to the desired standard is surveyed using the $\mathbf{t}_i$ plots and the Q plot, along all the positions that the moving window can take on the trans-section width. The confidence limits on the $\mathbf{t}_i$ charts and on the Q chart are not the same for the entire width, but are defined for each spatial window. In particular, the confidence limits of the Q chart take into account the different variability between valleys and edges: higher limits are set for the valleys because of a more marked shape variability; lower limits are set for the edges, in which the shape is expected to be only slightly variable.

Figure 13a shows (full thin lines) the five light intensity (trans-section) profiles on an edge segment that conforms to the required quality standards. These profiles are compared to the average light intensity profile on the reference segments (broken thick line). All test profiles are aligned along the left-hand rising branch of the reference profile. According to a rough visual inspection of these profiles, the inspected segment seems to conform to the quality standards. However, it is not completely clear whether the shape differences on the center and on the borders of the trans-section profiles are to be regarded as "regular" or not. A more rigorous and automatic monitoring of the edge shape can be done by analyzing the $\mathbf{t}_1$ plot (Figure 13b) and the Q plot (Figure 13c). It can be seen that no violations of the confidence limits are detected along the trans-section width (which is scanned by the moving window). Hence, Figures 13b,c unambiguously designate the tested edge as an on-quality one.

A similar comparison is presented in Figure 13d for a nonconforming segment. Again, although the analysis of Figure 13d is not completely satisfactory to assess the quality of the trans-section profile, the analysis of the $\mathbf{t}_1$ plot (Figure 13e) and of the Q plot (Figure 13f) unambiguously points to the locations (along the trans-section width) where the

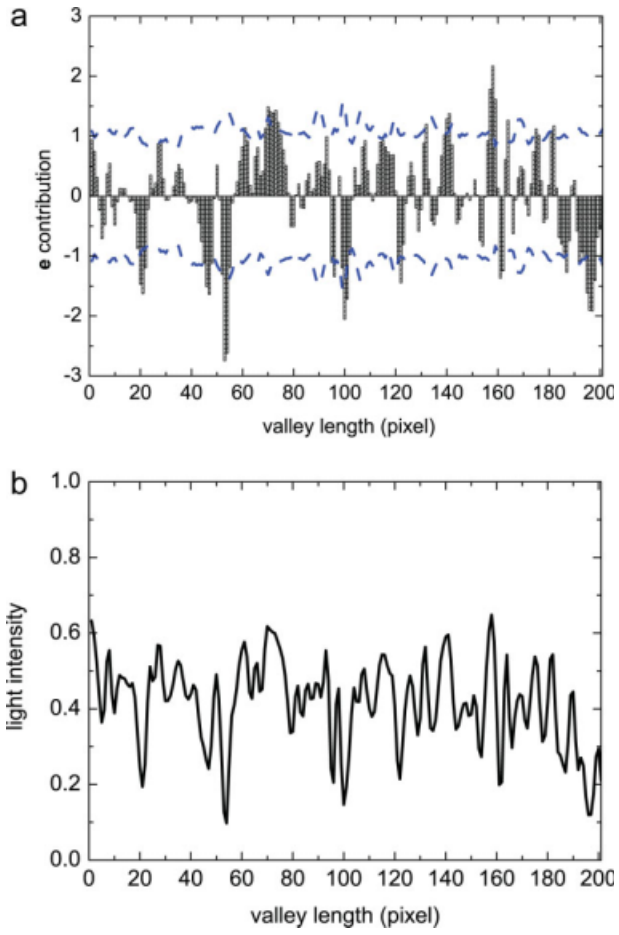

Figure 12. Surface roughness monitoring: identification of defects on a valley from the analysis of (a) the contributions to the residuals and (b) the profile of light intensity along the valley length.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

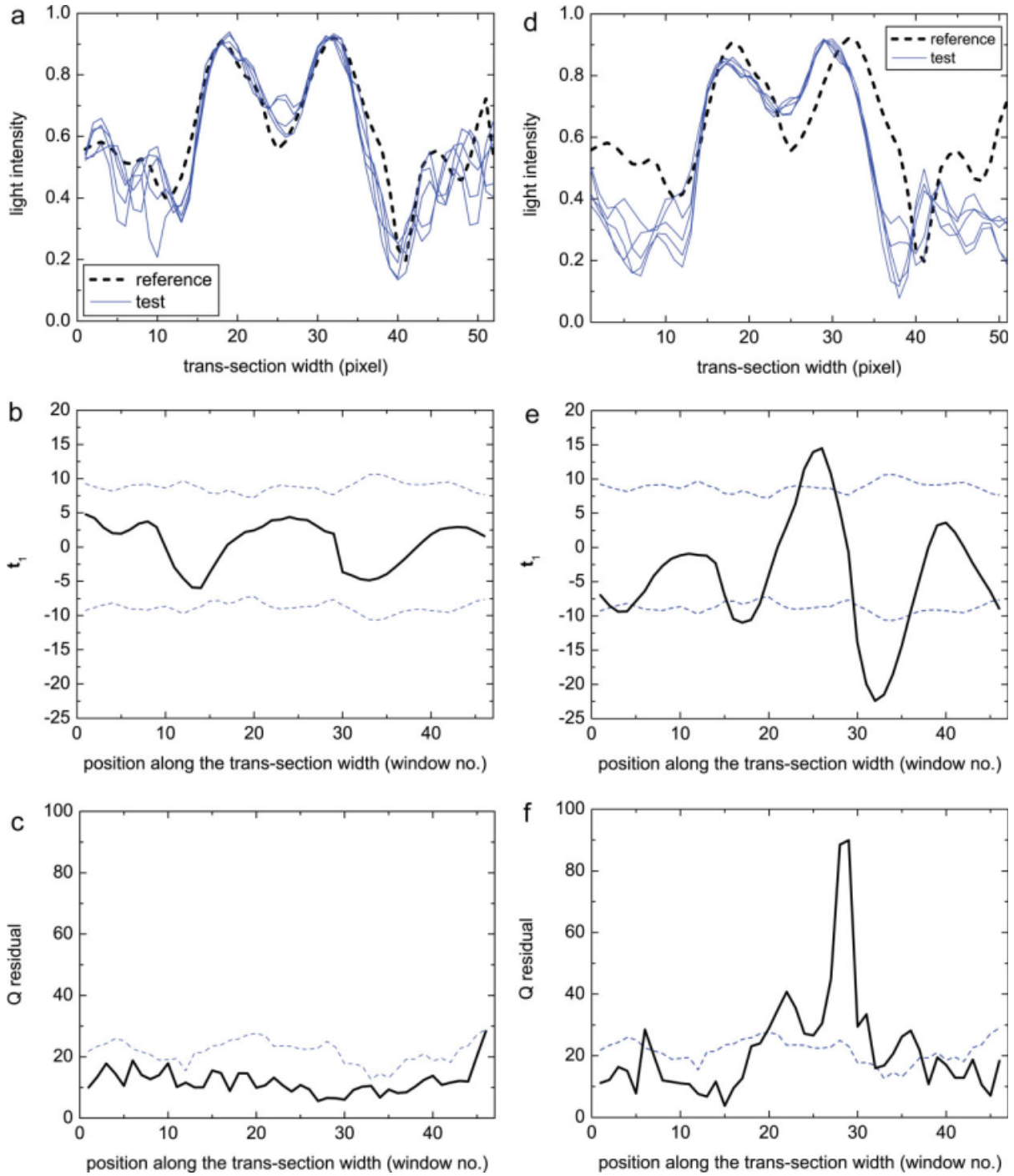

Figure 13. Edge shape monitoring for an on-quality segment (left column plots) and for an off-quality segment (right column plots).

(a) light intensity profiles on one segment of a test edge conforming to the required quality standards; (b) first score as a function of the moving window number for the same conforming segment; (c) Q statistic as a function of the moving window location for the same conforming segment; (d) light intensity profiles on one segment of a test edge not conforming to the required standards; (e) first score as a function of the moving window location for the same nonconforming segment; (f) Q statistic as a function of the moving window number for the same nonconforming segment. The dashed lines represent the confidence limits. [Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

nonconformity is present. Note that this test segment is categorized as an "on quality" one in the valleys, although Figure 13d shows that there is a significant difference between the borders of the test and those of the reference profiles. However, the spatial moving window approach effectively allows to identify the difference in the roughness structure between edges (smoother structure) and valleys (coarser structure).

Conclusions

One of the most important step in the integrated circuit fab is photolithography, because of its economical and operational impact on the manufacturing scheduling. The greatest part of the monitoring efforts in photolithography is focused on the measurement of the most important physical parameters of a photolithographed device, such as CD, LER, or SWA and the common inspection tools are optical instruments.

An image, however, retains information that largely exceed the mere metrology, and which are useful to identify the complex nature of the manufactured product quality. Through image analysis, and in particular advanced image processing techniques, it is possible to access without human intervention several meaningful clues that help to better understand the manufacturing process, to identify the critical situations (for the product quality and the process progress) and to counteract the problems.

In particular, in this paper, a monitoring system for the after-development quality evaluation was introduced. The approach, based on a combination of multiresolution and multivariate statistics, can deal with the hidden characteristics of a device, performing all the tasks of the instrumental sensors, but grasping additionally features of the surface product, which are commonly inaccessible. Being quality a multivariable property in nature, multivariate statistical techniques were exploited to extract the information embedded in the image. Although the signals were corrupted by noise and alterations, which affects different scales of resolutions, only the relevant scales were considered through a multiscale treatment by means of the wavelet decomposition. The result was a multiscale and multivariate monitoring framework that inspects the quality of the photolithographed device through the analysis of a SEM image.

The proposed monitoring system was shown to be an effective tool for the survey of the product. By adding new features for the measurement of the critical physical parameters, the image analysis system performs a full scanning of the surface, identifying and localizing the defects and anomalies, and detecting in advance the drifts of the process.

Acknowledgements

P. F. gratefully acknowledges "Fondazione Ing. Aldo Gini" and Sirca S.p.A. for financial support.

Notation

$\mathbf{e}$ = residual of a test sample
Δn_{segm} = segment width (pixel)
Δn_{mw} = size of the moving window (pixel)
i_M = filtered image at the M-th scale of decomposition (squared pixel)
j = counter
m = decomposition scale
M = selected decomposition level
M_1 = decomposition level selected for the image de-noising
n = counter
n_{el} = size of the edge length (pixel)
n_{iw}^{E} = size of the image width for the edges
n_{iw}^{V} = size of the image width for the valleys
n_{levels} = number of selected topological levels
n_{tsw} = size of the trans-section width (pixel)
N_{image} = number of images of edge segments
Q = square prediction error of the PCA model
$\Re$ = set of the real numbers
$\mathbf{s}$ = variable in the domain of pixel space (squared pixel)
$S_{M,n}$ = approximation coefficient at the M-th scale of decomposition
$\mathbf{t}_1$ = score vector of the first principal component
$\mathbf{t}_i$ = score vector of the i-th principal component
T^2 = Hotelling statistics
$T_{m,n}$ = detail coefficient at the m-th scale of decomposition

Greek letters

ϕ = father wavelet
ψ = mother wavelet

Acronyms

2-D = bi-dimesional
3-D = three-dimensional
CD = critical dimension
CD-SEM = tool for the measurement of the CD through a SEM
IC = integrated circuit
LER = line edge roughness
PCA = principal component analysis
SEM = scanning electron microscopy (or microscope)
SWA = side wall angle
UV = ultra violet

Literature Cited

1. Waits CM, Morgan B, Kastantin M, Ghodssi R. Microfabrication of 3D silicon MEMS structures using grey-scale lithography and deep reactive ion etching. *Sens Actuators A*. 2005;119: 245–253.
2. Blais P, Micheals M, Helbert J, ed. Issues and trends affecting lithography tool selection strategy. In: Helbert, JH., ed. *Handbook of VLSI Microlithography. Second edition: Principles, Technology and Application.* Park Ridge, New Jersey: Noyes Publications, 2001; 1–73.
3. Guldi RL. Inline defect reduction from a historical perspective and its implication for future integrated circuits manufacturing. *IEEE Trans Semicond Manuf*. 2004;17:629–639.
4. Yaakobovitz B, Cohen Y, Tsur Y. Line edge roughness detection using deep UV light scatterometry. *Microelectron Eng*. 2007;84: 619–625.
5. Edgar TF, Butler SW, Campbell WJ, Pfeiffer C, Bode C, Hwang SB, Balakrishnan KS, Hahn J. Automatic control in microelectronics manufacturing: practices, challenges, and possibilities. *Automatica*. 2000;36:1567–1603.
6. Yue HH, Qin SJ, Markle RJ, Nauert C, Gatto M. Fault detection of plasma etcher using optical emission spectra. *IEEE Trans Semicond Manufact*. 2000;13:374–385.
7. Waldo W. Techniques and tools for optical lithography. In: Helbert JH., ed. *Handbook of VLSI Microlithography. Second edition: Principles, Technology and Application*. Park Ridge, New Jersey: Noyes Publications, 2001;472–643.
8. Rao AR. Future directions in industrial machine vision: a case study of semiconductor manufacturing applications. *Image Vis Comput*. 1996;14:3–19.
9. Lee F. Lithography process monitoring and defect detection. In: Helbert, JH., ed. *Handbook of VLSI Microlithography. Second edition. Principles, Technology and Application*. Park Ridge, New Jersey: Noyes Publications, 2001;327–381.
10. Knight S, Dixon R, Jones RL, Lin EK, Orji NG, Silver R, Villarrubia JS, Vladár AE, Wu W. Advanced metrology needs for nanoelectric lithography. *C R Physique*. 2006;7:931–941.
11. Constantoudis V, Patsis GP, Tserepi A, Gogolides E. Quantification of line-edge roughness of photoresist. II. Scaling and fractal analysis and the best roughness descriptors. *J Vac Sci Technol B*. 2003;21: 1019–1026.
12. El Chemali C, Freudenberg J, Hankinson M, Bendik JJ. Run-to-run critical dimension and sidewall angle lithography control using the PROLITH simulator. *IEEE Trans Semicond Manuf*. 2004;17:388–401.

13. Sun, W, Mukherjee R, Strove P, Romagnoli JA, Palazoglu A. A multi-resolution approach for line-edge roughness detection. *Microelectronic Engineering* 2009;86:340–351.
14. Jackson JE. *A User's Guide to Principal Components*. New York: Wiley, 1991.
15. Kourti T, MacGregor JF. Process analysis, monitoring and diagnosis, using multivariate projection methods. *Chemom Intell Lab Sys*. 1995;28:3–21.
16. Helbert JN, Daou T. Resist technology - Design, processing and applications. In: Helbert, JH., ed. *Handbook of VLSI Microlithography. Second edition: Principles, Technology and Application*. Park Ridge, New Jersey: Noyes Publications, 2001;74–314.
17. Kosanovich KA, Piovoso MJ. PCA of wavelet transformed process data for monitoring. *Intell Data Anal*. 1997;1:85–99.
18. Addison PS. *The Illustrated Wavelet Transform Handbook*. London, UK: IOP Publishing, 2002.
19. Salari E, Ling Z. Texture segmentation using hierarchical wavelet decomposition. *Pattern Rec*. 1995;28:1818–1824.
20. Ruttiman UE, Unser M, Rawlings RR, Rio D, Ramsey NF, Mattay VS, Hommer DW, Frank JA, Weinberger DR. Statistical analysis of functional MRI data in the wavelet domain. *IEEE Trans Med Imag*. 1998;17:142–154.
21. Patsis GP, Constantoudis V, Tserepi A, Gogolides E, Grozevb G. Quantification of line-edge roughness of photoresists. I. A comparison between off-line and on-line analysis of top-down scanning electron microscopy images. *J Vac Sci Technol B*. 2003;21:1008–1018.
22. Geladi P, Grahn H. *Multivariate Image Analysis*. New York: Wiley, 1996.
23. Wold S, Geladi P, Esbensen K, Öhman J. Multy-way principal components and PLS-analysis. *J Chemom*. 1987;1:41–56.
24. Bharati MH, Liu JJ, MacGregor JF. Image texture analysis: methods and comparison. *Chemom Intell Lab Sys*. 2004;72:57–71.
25. Nomikos P, MacGregor JF. Monitoring batch processes using multiway principal component analysis. *AIChE J*. 1994;40:1361–1375.
26. Wise BM, Gallagher NB. The process chemometrics approach to process monitoring and fault detection. *J Process Control*. 1996;6: 329–348.
27. Nomikos P. Detection and diagnosis of abnormal batch operation based on multi-way principal component analysis. *ISA Trans*. 1996; 35:259–266.
28. Conlin AK, Martin EB, Morris AJ. Confidence limits for contribution pots. *J Chemom*. 2000;14:725–736.

Manuscript received July 4, 2008, and revision received Sept. 10, 2008.