INSPECTION IN SEMICONDUCTOR MANUFACTURING

Product quality is the cornerstone of a company's economic performance, and most firms attempt to improve product quality by increasing their production skills. Such "learning-by-doing" practices also reduce the unit cost of most product lines (1), whether the product is large, complex, expensive, and assembled, such as an airplane (2), or small, simple, cheap and nonassembled, such as a bar of soap. However, the difference in value added between product lines such as these will motivate their respective manufacturers to pursue contrasting strategies for quality improvement.

An avionics manufacturer makes relatively few airplanes that generate a substantial amount of revenue each, and the consequences of a product failure in the field are extremely serious. The avionics manufacturer will therefore include multiple inspection steps in the complex production process, in order to detect defects before the product is shipped. The cost of repairing a defect in an airplane during production is small when compared with the cost of rejecting the whole unit at the end of the production line. The avionics manufacturer is therefore inclined to fix a product during production and choose "uptime" as a productivity metric. The manufacturer would define uptime as the fraction of time spent making airplanes as opposed to the time spent repairing their defects. Uptime would increase as the manufacturer improves the quality of the product, demonstrating that quality efforts enhance productivity by reducing production costs.

When compared with an airplane, a bar of soap is a commodity product with low value added. A soap manufacturer therefore has to engage in high-volume production techniques to make a profit. Fortunately the consequences of a failure in the field are benign, and the cost of repairing a bar of soap may exceed the cost of producing it. A soap producer is there-

J. Webster (ed.), *Wiley Encyclopedia of Electrical and Electronics Engineering*. Copyright © 1999 John Wiley & Sons, Inc.

fore likely to have only one inspection step at the end of the process, at which the product is accepted or rejected. If the cost of inspection contributes significantly to the unit cost, the producer may even choose to control quality through statistical sampling techniques (3). The soap manufacturer will most likely adopt "yield" as a productivity metric and define it as the number of good bars of soap shipped divided by the total number of bars produced.

A soap maker can compensate for low profit margins by taking advantage of economies of scale. The manufacturer may invest in automated manufacturing equipment, which will dramatically increase the fixed costs associated with its production process. The soap maker can only amortize equipment by getting as close to 100% yield as soon as possible, and by remaining there consistently. The soap maker must engage in an iterative experimentation process: design an experiment, conduct the experiment, analyze the data, and feed the results back into the design of the next experiment. This process, called yield learning, repeats until all sources of yield loss are detected, identified and eliminated, or the cost of further experimentation exceeds the benefit of the knowledge gained. Ideally, the soap maker will predict yield numbers and adjust the yield learning rate in order to maximize profit over time. This process is known as yield management (4–7).

Yield learning and yield management are crucial in the semiconductor industry, where a production process comparable in sophistication to that of making an airplane may generate a product that commands the price of a bar of soap. Industry experts agree that over 80% of yield loss in an integrated circuit (IC)—the predominant product of the semiconductor industry—comes from process-induced defects (PIDs) such as particulates in process equipment. Yield learning in the semiconductor industry is therefore closely associated with defect reduction.

DEFECT REDUCTION CYCLE IN SEMICONDUCTOR MANUFACTURING

Figure 1 shows an algorithm for yield learning through defect reduction, which is used by the majority of semiconductor manufacturers (8). It begins with a yield model that attempts to predict the effect of PIDs on chip yield and to allocate defect budgets to semiconductor process equipment (9). Chip yield, which is defined as the ratio of good chips shipped versus the total number of chips manufactured, depends upon IC design and IC feature sizes, as well as the size and nature of a defect. It must be noted that all detected defects do not cause

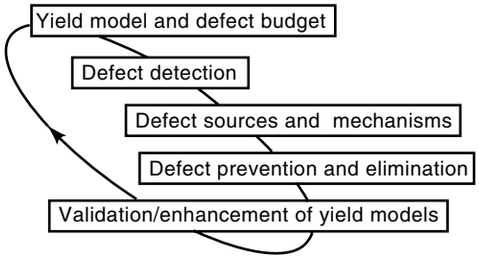

Figure 1. An algorithm for yield learning through defect reduction used by the majority of semiconductor manufacturers. The allocated defect budget is iteratively revised through cycles of detection, sourcing, prevention and elimination, to achieve higher yield.

electrical failures. The ratio of defects that cause failures to the total defects detected is termed as the "kill ratio." It is convenient to establish an empirical kill ratio at each inspection and to incorporate the empirical kill ratios of various defect classes into the yield model. This enables the yield model to convert the number of detected defects into an estimated number of anticipated electrical faults.

Defect inspection is the next step in the defect reduction process. Defect inspection procedures can broadly be classified as off-line and in-line, based on whether processing is interrupted while the products are inspected. Typically in-line inspection is nondestructive, whereas off-line inspection could potentially alter the product characteristics and render the tested specimen unusable. This classification also depends on the product specification that is being interrogated. While checking dimensions or for the presence and proper assembly of a single component, one can adopt in-line inspection, whereas checking for material flaws involves cross-sectioning a portion to examine microstructure. Nondestructive inspection techniques can be further classified as optical or nonoptical. The former category includes all techniques that utilize light as the sensing medium while the latter includes thermal, electron beam, ion beam, ultrasonic, radiographic, and electromagnetic approaches. Reference 10 provides an excellent overview of a wide variety of optical techniques. An exhaustive treatment of automated visual inspection can be found in Refs. 11 and 12. Several publications offering a historical perspective of inspection include Refs. 13 to 19.

A modern semiconductor process may consist of a sequence of more than 500 intricate process steps, which are executed on equipment that costs between $1 million and $10 million per unit. When one of these process steps drifts out of acceptable limits, it generates defects that may affect the downstream process. By detecting such drifts rapidly, the source of the problem can be identified and corrected before substantial wastage is incurred. Automated inspection tools must therefore be designed to fulfill the need of real-time feedback, and their information output must exhibit a format that is readily transferable. Process engineers, the users of inspection information, can subsequently control and improve individual process steps more rapidly, which contributes substantially to the improvement of the process as a whole (20).

The semiconductor industry relies heavily on optical techniques for defect detection and on nonoptical techniques for identifying defect sources and mechanisms, the next step in the defect reduction cycle. Defect sourcing generally begins by redetecting and reviewing the defect on a more sensitive tool. Machine vision software subsequently classifies the defect and converts the defect data into an easily transferable format. The knowledge gained through defect classification facilitates the identification of the source of the defect, and aids the design of an experiment that contributes to the eventual removal of its kind. The impact of eliminating a defect type is fed back into the yield model, effectively closing the defect reduction cycle. The cycle is repeated until the costs of experimentation exceed the benefits of yield learning (7).

The economic cost and benefit of yield learning and yield management include the opportunity cost of time, which affects the semiconductor industry significantly (7). The exponential relationship between defect density (defects per cm^2) and yield subject the industry to radical experience curves, which leading-edge manufacturers can exploit and convert into economies of scale (6). The price of an integrated circuit

has been known to deteriorate exponentially once a leading-edge manufacturer goes into volume production, adversely affecting the profitability of all followers (21–23).

Innovation drives the semiconductor industry as much as time to market. Moore's law, the mantra of the semiconductor industry, states that the transistor density on a manufactured chip will double every 18 months, and semiconductor manufacturers have teamed up with tool and materials suppliers to make that happen. Ever since Dr. Gordon Moore made his observation in 1965, the industry has produced goods of increasingly higher quality at successively lower cost. The computing power available per dollar has doubled every 12 to 18 months as a result. Thus the exponential growth of the computing sector over the past 30 years can primarily be attributed to continuous innovation in the semiconductor industry. Examples of such innovation include shrinking line widths, accelerating yield learning rates, improving overall equipment effectiveness, and augmenting silicon wafer sizes (24).

Silicon wafers are the basic unit of material in the semiconductor industry. They move from equipment to equipment during the production process, where they encounter processing techniques such as crystal growth (which forms the wafers), ion implantation, film deposition, lithography, etch, metal deposition, and polish. They are either subjected to processing tools individually or in lots. After a few weeks from the first step, the completed product wafer yields a batch of several tens to hundreds of chips.

Each process step adds value to the wafers, while simultaneously increasing the cost of the wafers and the economic consequences of failure. By the time a wafer reaches the end of the line, it may have cost as much as $3000 to produce. Incurring defects near the end of the process can therefore amount to significant loss of revenue to a manufacturer.

A chipmaker can also ill afford a failure in the field, because an integrated circuit may become a critical part of a critical system. For example, a defective chip in the navigation system of an airplane can cause the airplane to crash. A chipmaker therefore has a strong incentive to proactively inspect the product during the production process, just like an avionics manufacturer does. However, contrary to an avionics manufacturer, a chipmaker has no incentive to repair a chip during production; the unit price of an integrated circuit is simply too low for that.

In summary, the complexity of the semiconductor process environment prescribes a combination of sophisticated inspection tools and advanced process control methodologies for every successful defect reduction strategy (3). This article describes some of the technologies and methodologies needed for implementing such a strategy. The article concentrates on the inspection and analysis of patterned wafers, which is arguably the most formidable challenge in the defect reduction cycle.

INSPECTION IN THE IC MANUFACTURING PROCESS LIFE CYCLE

The first challenge of creating a complementary-metal-oxide-semiconductor (CMOS) integrated circuit is physically isolating and electrically insulating the individual devices from each other. In modern CMOS processes this is achieved by a technology called shallow trench isolation (STI). As the name suggests, trenches are etched into silicon and filled with the insulating field oxide shown in Fig. 2. Ion implantation into the field area prior to filling the trenches increases the electrical insulation of the devices. A very useful inspection, which can detect possible adverse effects of plasma etching on silicon bulk material, may occur at that point of the process. Another useful inspection, which covers all steps of the STI module, could occur after the trenches are filled.

The active area of the devices is located in all regions where the field oxide is absent. A series of ion implantation steps into the active area determine the threshold voltages of the devices. Upon completion of the active area implants, a gate oxide is grown and polysilicon is deposited. Etching the polysilicon defines the gate, source, and drain regions because the gate oxide remains where the polysilicon is not removed. High-dose, low-energy ion implants, followed by anneals, subsequently activate the source and drain regions of the devices.

An interlayer dielectric is deposited on top of the devices and planarized by chemical-mechanical polishing (CMP). Contacts are etched into the dielectric to gain access to the polysilicon line, as well as the source and drain regions. The aspect ratios of these contacts frequently exceed 10 to 1, which makes defect detection at the bottom of the contacts exceedingly difficult. Inspection of the contacts is also complicated by the fact that the bottom of polysilicon contacts reside about 300 nm above those of source/drain contacts, causing one or the other level to be out of focus.

In most modern semiconductor processes, contacts are plugged up with tungsten prior to the deposition of the first metal layer, which can also consist of tungsten or a complicated sandwich structure primarily composed of aluminum. After it is patterned and etched, the first metal layer typically serves as a local interconnect mechanism that links devices separated by only a few micrometers.

Multiple layers of metal with low resistivity connect more distant devices. Figure 2 only shows one such layer, but more than five have been known to exist in real semiconductor processes. In the upcoming 180 nanometer technology generation copper is the likely metal of choice for these upper metal layers. Materials with low dielectric constant will separate the copper lines to prevent cross talk. Since copper is difficult to plasma etch, the global interconnect lines are likely to be fabricated with dual damascene technology, where the vias that connect layers and canals are etched into dielectrics before metal deposition. Copper fills the vias and the canals. Excess copper is polished off using CMP techniques.

Two key inspections are likely to occur in dual damascene technology. The post-canal-etch inspection is intended to find subsurface defects, whereas the post-CMP inspection detects a variety of failure modes associated with chemical mechanical polishing (25).

Once global interconnect layers have been fabricated the completed CMOS-integrated circuit is passivated with a protective dielectric. Openings for pads are etched into the protective layer, and pads are fabricated. Contact pads are etched. Solder bumps typically connect wires to the pads.

Process development practices may vary from company to company, but all process development teams face the same physical challenges. In the early stages of research, engineers focus on solving fundamental process problems, relying primarily on computer models, on off-line metrology, and on electrical data from microelectronic test structures. Defect inspection becomes increasingly important once actual IC structures are introduced into the process development effort, when in-

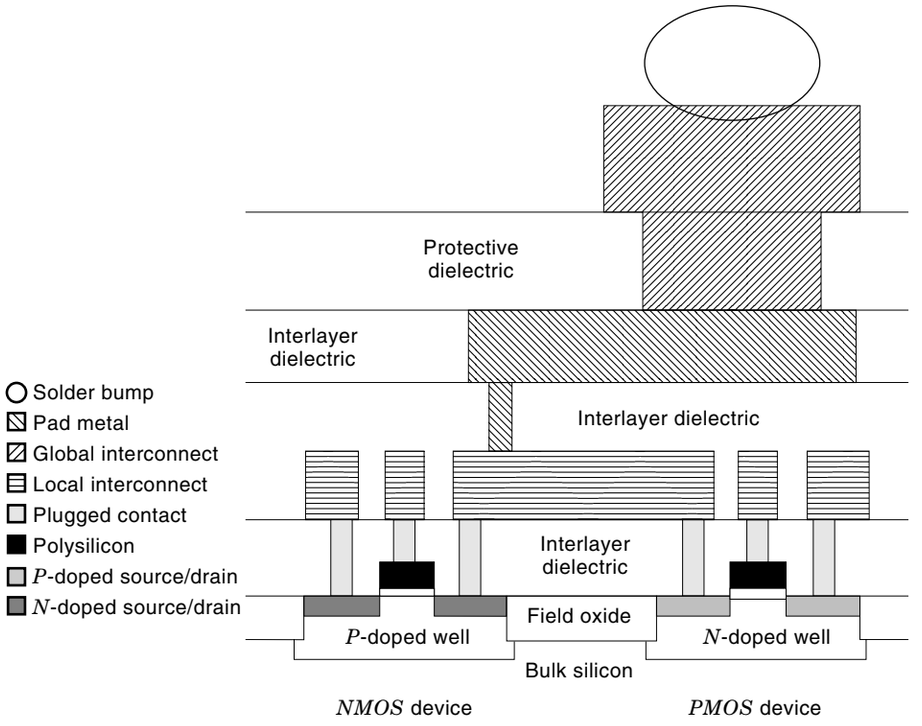

Figure 2. A cross section of a complementary-metal-oxide-semiconductor (CMOS) structure representing the materials employed and shapes fabricated at various process steps in semiconductor manufacturing. This diagram has been oversimplified to enhance its instructive value. It illustrates that semiconductor devices primarily consist of superposition of patterned and etched thin films.

spection tools detect and identify defects that could be the source of electrical faults. The requirement for these inspection tools changes as the process matures and number of observable defects decreases (26).

The 1997 edition of the Semiconductor Industry Association's *National Technology Roadmap for Semiconductors* (NTRS-97) has defined the following generic phases of the semiconductor process life cycle as they pertain to the evolving requirements of defect inspection (27,28).

Process Research and Development (PR&D). Relatively low production rates and yields, experimental development of process parameters, detailed examination of defects and identification of defects characterize the process R&D phase. Systematic faults and highly clustered defects dominate. Defects abound, so defect detection tools and methods require relatively low throughput.

Yield Ramp (YR). During the YR phase, the yield of a node-technology-driving product moves from approximately 20% to 80%. This phase is characterized by a number of defects on the product wafers that allows engineers to conduct statistically significant experiments (29) using product wafers at relatively low cost. The defect density is much lower than in the PR&D phase, which increases the throughput requirements of defect detection tools and methods. Random defects start to dominate as defect density decreases.

Volume Production (VP). Volume production represents the final stages of the life cycle of a semiconductor process, in which effectively no further tuning of the process control parameters is attempted. The objective of using defect detection tools in this phase is to identify process excursions as rapidly as possible, which requires tools and methods with very high throughput. Since the process is well seasoned by this stage, the problems are frequently catastrophic and could involve shutting down the line when defects are detected.

Within the semiconductor manufacturing industry, the scope of inspection and metrology is vast. Inspection of the photomasks, reticles, and pellicles used in lithography; inspection of bare wafers; and inspection of patterned wafers are some of the most important inspection steps. Subtle differences in the techniques and their end-application such as measurement of critical dimensions, film thickness measurement, and measurement of overlay of the current layer with respect to the previous layer are viewed as metrology tasks. In the interests of brevity and clarity, the discussion in this article will be limited to detection, characterization and root cause analysis of patterned wafer defects alone.

Patterned wafer defects consist primarily of particles of contamination on or near the surface, scratches, missing patterns, and area defects. An image of a patterned wafer can be seen in Fig. 3. The variation in film thickness across the wafer causes the multicolored reflection of the incident white light observed in the figure.

One critical requirement of rapid yield learning is the ability to identify, source and eliminate defect mechanisms as soon as they develop. During the 1970s and 1980s, most integrated circuit fabrication facilities (fabs) were able to perform the defect review and analysis function with manual microscopes in the fab and relied on electrical test at the end of the line to catch the remaining problems. However, starting in the late 1980s high-speed automatic inspection and review tools started becoming available, allowing the fabs to shorten the feedback loop for identifying, sourcing and eliminating the defect mechanisms as they occurred. The yield ramps of the 1990s would have been impossible without the advent of these automatic inspection and review tools. Projecting this trend forward, the experts in the field envision further automation of the

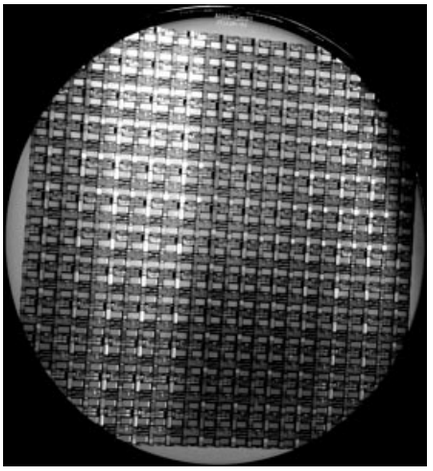

Figure 3. Image of a patterned wafer. Each square outline represents a chip. The multicolored reflection is a result of varying film thickness and surface roughness across the wafer.

entire yield learning and yield management process for the fabs of the future including defect inspection, review, analysis, source identification, and source elimination (8).

Although the defect metrology tool deployment strategy for a state of the art fab of the late 1990s with 0.25 μm process technology varies from company to company, there are some common elements found in most fabs. Figure 4 graphically depicts the inspection, review and data management set up of a typical semiconductor fabrication line. For a logic product fabrication line running a 0.25 μm process flow with functional chip yield in the 20% to 80% range (yield ramp phase), most fabs do some sort of automatic defect inspection and review at 12 to 16 critical mask levels of the 20 to 25 mask levels in the process flow. When a new process technology is introduced in the manufacturing fab (process development phase), the first few lots are inspected at each mask level, whereas a fab running at more than 80% functional yield (volume production phase) typically inspects the lots at fewer steps.

Automated Detection, Review, and Analysis

A typical automatic lot inspection, review and analysis process in the yield ramp phase consists of the following steps (8):

1. Automatic defect detection with a tool employing optical image processing or laser scattering technology; both these technologies are described in greater detail later. These tools are set up with a set of operating parameters selected by engineering personnel for all inspection process points for each product. Each set of tool parameters is called a recipe; the manufacturing technician inputs the right recipe for the lot at hand with the sampling criteria provided by engineering. Inspection samples vary from two to ten wafers per lot; typically 30% to 100% of all lots are tagged for inspection. Most lots exhibiting normal results and are returned to the process flow after automatic defect inspection. All excursion lots go for review as detailed in steps 2 and 3 below. Sometimes lots that are in-control lots are also reviewed.
2. On-line or off-line spatial signature analysis (SSA) (See Fig. 5) and automatic defect classification (ADC) with an optical or scanning electron microscope (SEM) tool; a defect library of commonly occurring signatures and defects is set up by the engineering personnel for each process level and is used as reference during lot testing. Defect images are captured, defects are classified, and images are stored for future reference.

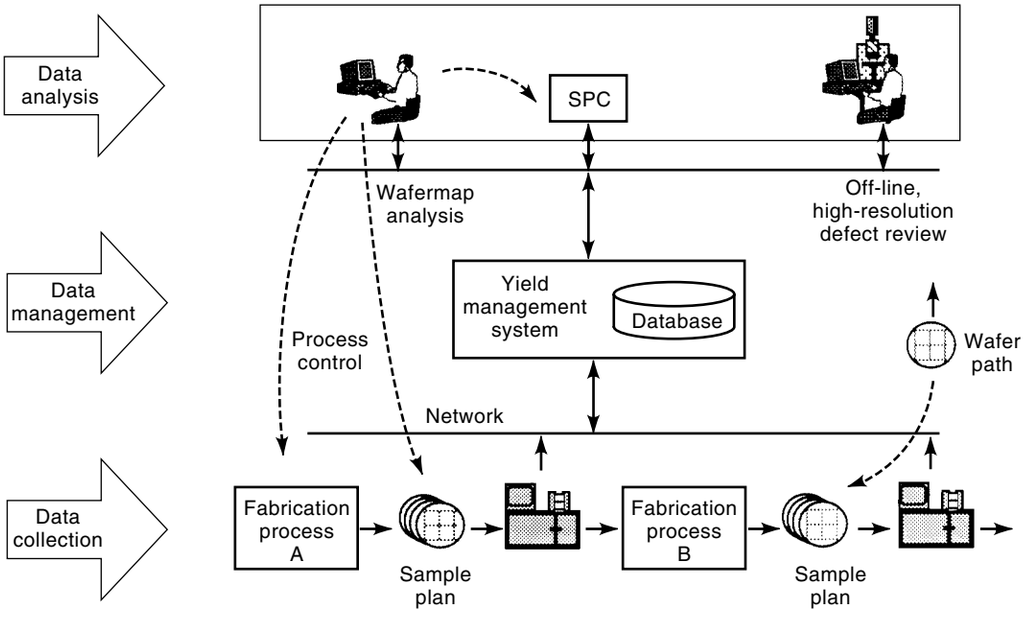

Figure 4. The flow of wafers and management of data generated during fabrication and inspection of ICs is schematically described in this figure.

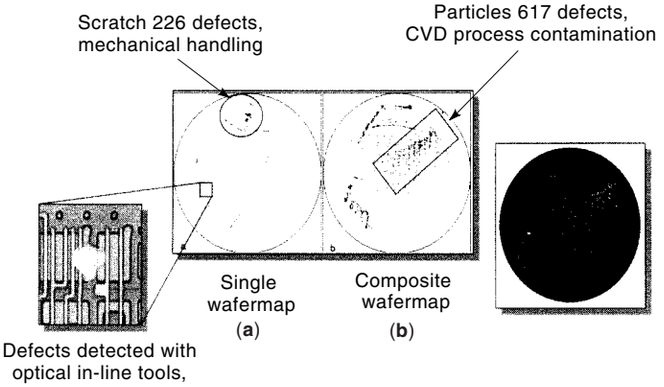

Figure 5. A spatial signature is defined as a unique distribution of wafer defects originating from a single manufacturing problem. (a) Single wafer containing scratch signatures; (b) stack of wafers superimposed highlighting a subtle systematic particle contamination problem.

3. Following defect review on a SEM, energy dispersive X-ray spectroscopy (EDX or EDS) is performed on a small number of defects with unknown material composition. The EDS analysis results are used to assist in identifying the source of the defects and the responsible process engineer is notified. Depending on the severity and the frequency of occurrence of a given defect class, the process engineer either proceeds with the corrective action or alerts relevant process personnel to be on the lookout for the particular defect class on future lots. This step of the defect inspection and review sequence is the most labor intensive at this time and it needs to be automated soon to maintain productivity levels.

Other Issues

Materials-handling issues are very critical to ensuring reliable and consistent automation of inspection. Wafers are moved around the fab in wafer cassettes. All wafers are placed into slots arranged 25 to a cassette. The cassette itself is placed in a box that supports the bottom half. The lid of the box snaps onto the lower half and ensures minimal internal movement of the wafers. The box is designed to be airtight to prevent contamination of the wafers.

During inspection, the operator loads the cassette into a loading port. In most tools the loading port requires that the wafers are horizontal with the open half of the cassette facing the tool. A sensor quickly scans the cassette to determine which of the slots are occupied and relays this information to the inspection software. A wafer-handling robot extracts one of the wafers to be inspected by sliding a paddle beneath the wafer, lifting it slightly, and retracting the paddle. The wafer is then transferred onto the motion stage of the inspection tool.

Usually alignment of the wafer is the first operation performed. This involves locating the center of the wafer and the notch to determine the coordinate transformations with respect to the detector system. This subsystem rotates the wafer under a fixed camera which is located approximately tangential to the wafer circumference, looking down. The off-center shift of the wafer with respect to the rotation axis, and the wafer notch can be determined by thresholding the sequence of images acquired by the camera during one rotation. In some tools a prealigner performs this task; it rotates and translates the wafer to a predetermined orientation before loading. Since demands on the coordinate accuracy of the overall system are very high, this alignment task has an extremely low error margin. Additionally automated inspection requires every wafer to be oriented identically with respect to the illumination source. Incorrect wafer tilt can generate meaningless results, thereby confusing the operator and wasting other resources.

OPTICAL IMAGING TECHNOLOGY

Optical imaging techniques for semiconductor wafer analysis fall into roughly three main categories: (1) in-line microscopy for rapid, whole wafer defect detection; (2) off-line microscopy for defect review and failure analysis; and (3) microscopy techniques for critical dimension (CD) and overlay metrology.

In-line microscopy must keep up with the flow of manufacturing. State-of-the-art in-line imaging systems require from 2 to 20 min to scan a 200 mm wafer. At a defect sensitivity of 0.25 μm, a throughput of approximately 2 GHz is required, namely 2×10^9 image pixels must be captured, moved through the image processing system, and reduced to a small set of descriptive features, such as defect location and size, per second. These systems must therefore rely on high-speed and reliable imaging technology. Optical microscopy using charge-coupled devices (CCD) are the systems of choice for in-line defect detection today with CCDs configured in line-scan or time-delay integration (TDI) format. Off-line review typically involves re-sampling a small fraction of the wafers inspected in-line. Off-line review tools therefore do not require the same high throughout but do depend on higher image fidelity so that human-level decisions can be made regarding defect type and source. Higher-quality images for failure analysis are obtained by using, for example, color CCD sensors, confocal microscopy, and scanning electron microscopy (SEM). The following section will consider these most common methods of optical and electron microscopy used in semiconductor manufacturing today for both in-line detection and off-line review.

Sensors for Image Acquisition

CCD sensors are used almost exclusively as imaging detectors for optical microscopy. A CCD sensing element is a MOS capacitor that converts light photons into electrons (30–32). These sensing capacitors are arrayed in linear and area configurations to provide line-scan and area imaging capabilities (33). The spectral response of the CCD is a function of the semiconductor material properties. Long wavelength light will tend to pass through the sensing region of the MOS element and be absorbed deep in the substrate below the active area of the sensor. Short wavelength light will tend to be absorbed in the passivation layer above the active area of the device. Therefore silicon CCD imaging devices are sensitive between 400 nm (blue) and 1000 nm (red), with peak sensitivity around 700 nm. The devices are operated by integrating the incident photon energy over a short period of time and then transferring the recorded charge out of the device as an analog signal.

Thermal effects complicate the operation of CCD devices. At room temperatures, electron-hole pairs of sufficient energy are created in these devices that fill the storage region and cause a contamination of the signal. Applying a positive voltage to the silicon gate, resulting in an initial electron deficiency prior to integration mitigates this thermal effect. This effect is realized in the CCD electronics. The user improves the signal-to-noise ratio (SNR) by controlling the integration of light on the sensor, namely the exposure to light between charge transfer cycles. The allowable time for integration is driven by the required image quality and by the application (34). A longer integration time requires that the inspection surface persist in the field of view of the imaging sensor, or pixel, for a longer period of time. This proves problematic in a high-speed, high-resolution imaging environment since these conditions together translate into short integration times. To accommodate high speed and high resolution, for in-line inspection, the dominant CCD sensor configuration is the TDI sensor (35).

The TDI is an area array of imaging elements that works as a line-scan integration device. Figure 6 shows how the CCD TDI operates. A linear region of the wafer surface is imaged onto the first line of the TDI sensor at time t_i. After a predetermined integration time, the wafer has shifted during continuous scan and a new line on the surface is imaged onto the first CCD line. Meanwhile the charge from the first sensor has been shifted to the second CCD line, at time t_{i+1}, and exposed to the same viewing region of the wafer as the previous line. This process continues down the array resulting in an n-fold integration of the imaged line, n being the number of integration lines available on the TDI sensor. The output of the TDI is a single line of integrated data with improved SNR, analogous to the output of a common line-scan CCD device. As the wafer passes under the imaging optics, a semiconductor device image is built up line by line.

The optical system used to image a semiconductor device is designed to enhance various detail of the wafer surface prior to analysis. Bright-field (BF) and dark-field (DF) imaging is used to improve contrast between different materials and topology respectively, such as due to absorption and scattering or due to surface texture. Figure 7 shows optical configurations for BF and DF imaging. In BF imaging, the illumination path is coincident with the viewing path. Contrast in a BF image is a strong function of light attenuation and reflection between differing materials. Contrast in DF imaging, on the other hand, is a result of indirect illumination such that none of the light rays will be collected by the imaging optic unless scattered by an edge or a textured surface.

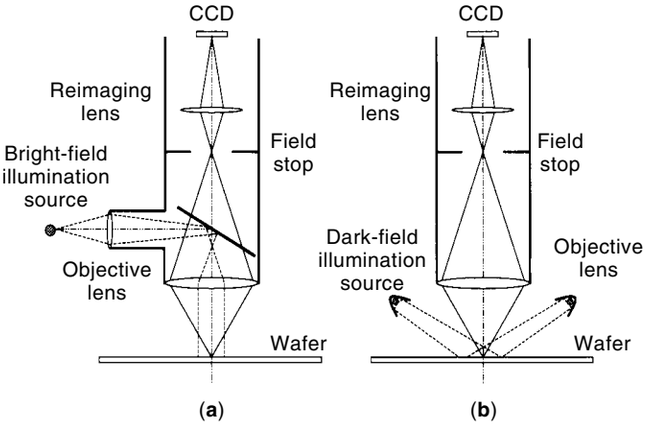

Figure 7. This representation of bright-field imaging (a) shows how the configuration allows collection of directly reflected light; the dark-field configuration shown in (b) collects only the scattered light.

Figure 8 shows an example of a BF (a) and DF (b) image of a finished semiconductor device at about 5 μm per pixel resolution. The device is passivated and has a large stain defect in the center of the field of view. Note in the BF image that the die pads contrast highly against the passivated surface but that the stain is difficult to discern. In the DF image, the stain, which is smooth relative to the device topology, is highly contrasted. The DF image also tends to highlight edge information, such as the perimeter of die pads. In practice, an imaging system may employ some level of both BF and DF illumination within the same system. The relative contribution would be established by the user to produce a high contrast image for a given inspection point and/or product. This tool setting would be maintained as a component of an inspection recipe.

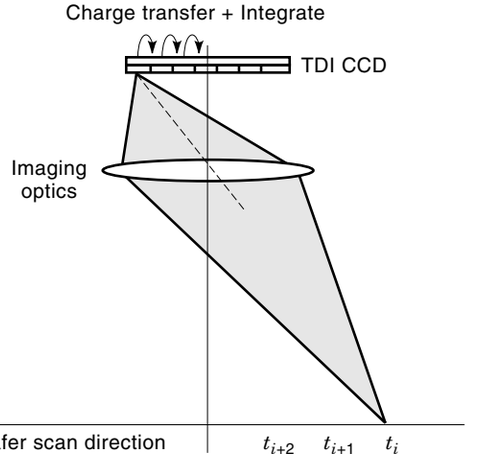

Figure 6. This schematic of a time-delay integration CCD device illustrates the mechanism for obtaining high-resolution images rapidly by moving the object relative to the camera.

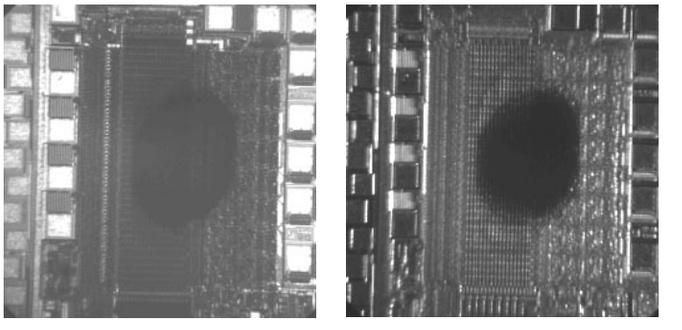

Figure 8. These images of a semiconductor device acquired at approximately 5 μm resolution of illustrate the difference between bright-field (a) and dark-field (b) imaging.

As the critical dimensions (CD) of semiconductor devices continue to shrink, defects that were too small to electrically interfere with a circuit are becoming problematic. A defect of roughly one-half to one-third this dimension can cause electrical failures through bridging at today's CD of 0.25 μm. Defects as small as 80 nm can cause electrical faults but are not visible by means of optical inspection (36). At this resolution, SEM imaging can provide high defect resolution and contrast, although electron/semiconductor interactions are vastly different from photon/semiconductor interactions (37). The primary role of SEM in today's fabrication facility is for off-line review, failure analysis, and CD and overlay metrology. Figure 9 shows the interaction of a primary electron beam with a material surface. The interaction products are composed of backscattered and secondary electrons, along with other byproducts such as Auger electrons and characteristic X rays (38). Standard SEM relies on a collection of secondary electrons with a photo-multiplier tube to build up an image while scanning the primary beam across the surface. These images are highly resolved and have a large depth of field compared to optical microscopy. Secondary electrons are generated within the top 1 nm to 10 nm of the sample surface and therefore image surface topology and structure. Backscatter electrons provide contrast between material properties since the scattering mechanism is a function of atomic number. Auger electrons and X rays are capable of differentiating material and chemical composition but require extended integration times and are therefore typically applied to small, localized areas, such as in identifying the composition of a particle defect.

SEM is very effective for revealing surface structure, defect morphology, and elemental composition, but current tools are slow and only now beginning to provide a high degree of automation. As the CD continues to shrink, advances in SEM are required that will result in the application of SEM to high-speed, in-line applications.

Resolution, Data Rates, and Throughput in Optical Imaging

Finding a 0.25 μm defect on a 200 mm diameter wafer is equivalent to searching for an object the size of a baseball on 58,000 acres of land (or one part in 5×10^{11}). To achieve specified detection sensitivity on an inspection tool, it is required that the surface be sampled at the appropriate resolution. At 0.25 μm resolution, Nyquist sampling theory requires a sample size of 0.125 μm (39,40). While this may be sufficient to detect a 0.25 μm defect, the ability to resolve it requires on the order of five times Nyquist, or about 50 nm per sample. This becomes especially critical when automating defect classification, as will be discussed in the following section.

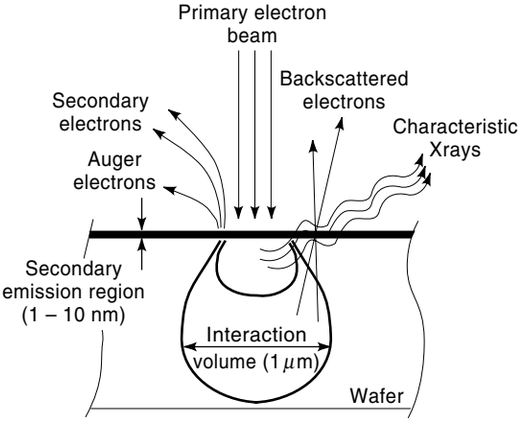

Figure 9. Depiction of typical interaction of the primary electron beam with a sample during SEM imaging.

As mentioned earlier, to maintain a throughput of 200 mm wafers at 3 wafers per hour requires a pixel processing speed of about 2 GHz. This was determined by dividing the wafer area by a Nyquist sample size of 0.125 μm (for a 0.25 μm sensitivity specification) and dividing by 3 wafers per hour. A series of images must be captured, processed, and reported at this sampling rate. This is accomplished by scanning the wafer across the rows of die and subtracting one die from its neighbor to locate subtle differences. While this process will be detailed in the next section, the following generic algorithm describes the required steps for defect detection:

1. An image is captured from a test region of a die.
2. The wafer stepper scans to the same location on a neighboring die, and a second image is captured.
3. The two images are aligned.
4. The images are subtracted.
5. Small alignment residuals and texture anomalies are filtered and removed.
6. The defect location and size features from the resulting mask are extracted.
7. Defect information is logged in the electronic wafermap file for later reference.
8. The stepper moves to the next die and the process is repeated.

To maintain the throughput suggested above, it is required that several of these steps be completed in parallel. For this reason parallel, pipeline image processing is a leading technology in this area. Pipeline image processing allows for a serial stream of functions to be processed in parallel. As a simple example, an algorithm consisting of image transfer (from camera to memory buffer), image subtraction, and image filtering can be processed with each functional block running in parallel. There will be a phase delay, or time lag, between the first image in and the resulting output, but this will be constant. As digital signal processing (DSP) devices and field-programmable gate arrays (FPGAs) continue to mature, these devices are also being used more often in parallel configurations (e.g., multiple DSPs), and in conjunction with pipelined architectures. The reconfigurability of DSPs and FPGAs coupled with their potential for low cost makes them attractive alternatives to the more structured architecture of pipeline processors.

LASER-SCATTERING TECHNOLOGY

Measurement of optical scattering is a very powerful mechanism to detect microscopic deviations on extremely smooth surfaces at high speed and high sensitivity. The inherent roughness of bare wafer surfaces is on the order of hundreds of angstroms. Any defect on or near the surface of the wafer can be detected by scanning a coherent spot of light over the

entire wafer and monitoring the intensity of the scattered light in nonspecular directions (also called the dark field). This approach forms the core of some inspection systems available today.

Photon scattering can be defined as the process by which energy is removed from a light beam and reemitted with a change in direction, phase or wavelength. It is an instantaneous process; if there is a measurable delay between absorption and reemission, the mechanism may be termed luminescence. Lord Rayleigh (41) derived an equation, given below for unpolarized light, for scattering by particles smaller than the wavelength of the exciting light

$$\frac{I(\theta)}{I_i} = \frac{n\pi V^2(1+\cos^2\theta)(m-1)^2}{d^2\lambda^4}$$

where

$I(\theta)$ is the intensity of scattered light from an incident beam of wavelength λ and intensity I_i at a distance d,
n is the number of scattering particles,
V is the volume of the disturbing particle,
θ is the angle of scatter, and
m is the index of refraction of the scattering medium.

This theory holds true for very small particles only, but is valid for any refractive index. G. Mie extended this theory to be applicable to larger particles but his solution is valid only if the refractive index m is close to unity (42).

Polarized light rays can be considered as waves vibrating in planes orthogonal to the direction of the ray. As they are scattered from wafer surfaces or particles, particularly at certain angles, wave components in some direction are absorbed more than those in other directions. Light waves whose electric field vibrates in the plane perpendicular to the plane of propagation of the beam (plane formed by incident and reflected beam) are considered as s-polarized; those with an electric field vibrating in the plane parallel to the plane of propagation are considered as p-polarized. This property is particularly useful for discerning the type of scattering particle from out-of-plane scattering measurements (43–45).

Since the random height variations of the wafer surface are much smaller than the wavelength of the incident light, speckle does not affect measurements. In this domain the scatter signal is dominated by diffraction from surface topography. By measuring the intensity of the diffracted orders, the bidirectional reflectance distribution function (BRDF) can be approximated by the expression

$$\text{BRDF} = \frac{P_s}{\Omega \cdot P_i \cos\theta_s}$$

where

P_s is the light flux (watts) scattered,
Ω is the solid angle (steradians),
P_i is the light flux per unit illuminated surface area, and
θ_s is the scattering angle.

The cosine function in the BRDF calculation causes a maxima at an angle normal to the surface of the wafer as shown in Fig. 10. If a library of typical BRDF values at known scattering angles can be constructed for a wafer with a given film stack, the measurements can be accomplished quite simply. It then becomes possible to accept or reject a wafer on the basis of appropriate specifications and measured values (46).

Sensors for Scatter-Based Defect Detection

Scattering tools generally use a photomultiplier tube (PMT) as the sensor for detecting the scattered light. PMTs can be used to detect wavelengths from 180 nm to 1200 nm. A PMT is extremely sensitive, allowing the measurement of very low levels of light. It has a wide dynamic range, so it can also measure high levels of light. It is very fast, so rapid spectral events can be reliably monitored. PMTs are also quite robust. When properly cared for, a PMT will typically function for 10,000 to 100,000 h.

The photomultiplier has a light sensitive electrode called the photocathode formed of semiconductive material containing alkali metal. The active layer, which emits electrons when photons strike it, can be on the surface of the metallic support or in semitransparent form on the inner surface of the silica envelope. These electrons are accelerated by a series of electrodes, called dynodes, toward a collector (or anode). The electrons produce several secondary electrons each time they strike a dynode resulting in a multiplication of their number as they approach the anode. Figure 11 shows a photomultiplier tube. An external resistor chain connected to a stable power supply is used to produce the voltages which are applied to the dynodes.

Typically each electron which strikes a dynode will produce about four secondary electrons. This means that if one electron is released from the photocathode, a PMT with 10 dynodes will deliver 4^{10} electrons to the collector. This constitutes a pulse that typically lasts for about 5 ns, so the resulting anode current is of the order of 1 mA. This gain, of about one million, is critically dependent on the dynode voltages which necessitates a very stable power supply.

Two factors determine the performance of a PMT. First, the conversion efficiency of the photocathode (photocathode sensitivity) varies with the wavelength of the incident light. This relationship between photocathode sensitivity and wavelength is called the spectral response characteristic of the PMT. Second, a small amount of current flows in a photomultiplier tube even when the tube is operated in a completely dark state. This output current, called the anode dark current, and the resulting noise are critical factors in determining the detectivity of a photomultiplier tube. Dark current is greatly dependent on the supply voltage.

Photon counting is the best way to operate a photomultiplier tube to measure low-light levels. In photon counting mode, individual photons that strike the photocathode of the PMT are measured. Each photon event gives rise to a pulse, or a count, at the output. The number of pulses, or counts per second, is proportional to the light impinging upon the PMT. The tube is typically operated at a constant high voltage where the PMT is most sensitive. Because of noise from various sources in the tube and the electronics, the output of the PMT will contain pulses that are not proportional to the light input. The detection system must reject these spurious pulses. It does this with a discriminator that electronically "discriminates" a low-level noise signal from a higher-level

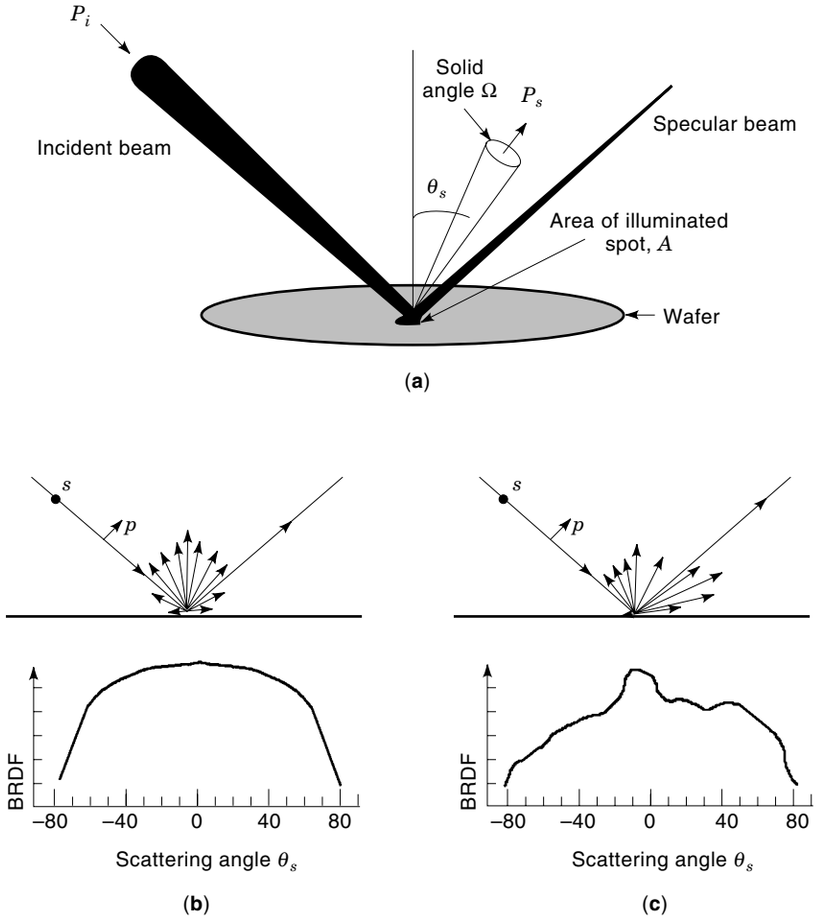

Figure 10. (a) This is an illustration of the optical arrangement and parameters to measure scattering cross section. (b) Typical angle-resolved BRDF measurements obtained from a clean surface and (c) from a contaminant particle on the surface.

signal from a photon event. As the number of photon events increase at higher light levels, it becomes difficult to differentiate between individual pulses and the photon counting detector becomes nonlinear. This response is also called saturation of the detector and usually occurs at 1 to 1.5 million counts per second. In that case it is customary to decrease the light level or switch to analog detection.

In analog detection, the high impedance at the anode and the stray capacitance effectively averages the pulses from the photocathode. The resultant photocurrent is converted to a voltage that is proportional to the intensity of the light striking the photocathode of the tube. The high voltage driving the tube may be varied to change the sensitivity of the PMT. Analog detection is usually used to measure medium-to-high light levels. At low light levels, analog noise may prohibit measurements of highest quality.

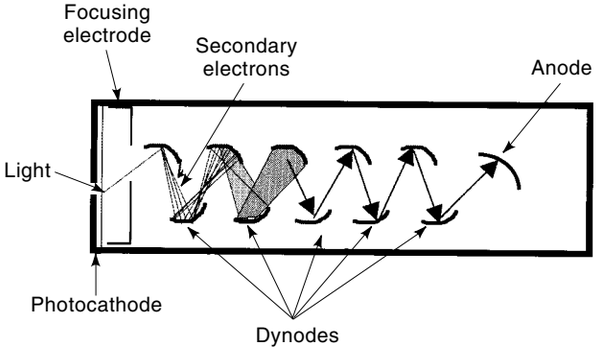

Figure 11. Schematic of a photomultiplier tube (PMT) illustrating the conversion of incident light detection into electric current.

MEASUREMENT OF OPTICAL SCATTER FROM CONTAMINANTS ON WAFERS

The key components of an optical scatter-based defect detection tool are a laser, focusing and aligning optics, and PMT (or other) detectors. Scattering is usually measured in the dark field where it is common to encounter low photon counts. The laser power, the sensitivity of the detectors, and their dynamic range drives the signal to noise ratio of the defect detection tool. Microroughness of the wafer surface creates a background scatter that is known as haze. This is acceptable and the tool should be able to tolerate it (47). It is well known that particles will scatter differently based on incident angle and polarization. The configuration of the detectors must eliminate all possible beam interactions with the scattered light. Also at different process steps the surface characteristics and the film stack will dictate the scattering behavior (48,49). Given a fixed configuration, each detection tool performs better on certain process steps than on others. For ex-

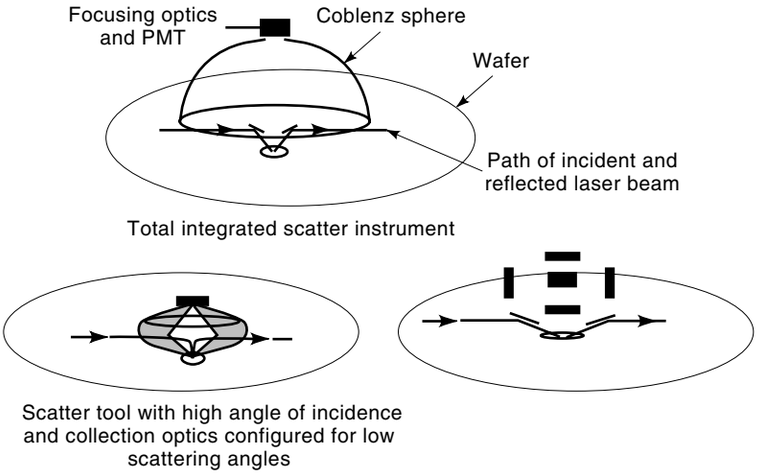

Figure 12. Three possible configurations of the source and detectors in scattering-based inspection tools.

ample, a tool that has a high incident angle illumination can be used to detect particles on smooth silicon and dielectric films, whereas a glazing angle illumination would be required to detect particles on rough or grainy surfaces. At the same time the tool should record minimal numbers of "nuisance" defects such as color variation, grain, and previous level defects.

The following are a few source-detector configurations:

1. Total integrated scatter (TIS) uses a device known as a Coblenz sphere to focus all the scattered light in the hemisphere above the particle on to a small spot which is then measured using a PMT.
2. Angle resolved scattering measurements involve positioning a detector at different discrete angles around the semicircle and measuring the intensity at each position. The ability to analyze the signal from different angles separately provides better discerning capacity of the type of particle or defect that is scattering the light.
3. Some tools collect light scattered from different arc segments of the hemisphere and focus them on separate PMT detectors. This can be thought of as an intermediate configuration to the above two.

These configurations are shown in Fig. 12.

Operational Details of a Scattering Tool

The laser beam is directed through a beam expander toward an oscillating mirror or prism and from there through a lens on to the surface of the wafer. The wafer is moved by the *X-Y* stage in a direction perpendicular to the scanning beam to cause the beam to sweep over the wafer in a series of adjacent scan lines. The incident light is scattered by particles, by surface defects, and by pattern features formed on the wafer. The scattered energy is collected by light collection optics and directed to a PMT for conversion into an electrical signal that corresponds to the detected intensity of the collected light. The aperture of the collection optics can reduce or increase the relative amounts of light from defects against pattern features. It also controls the solid angle of collection that determines the DSC. Knowing the incident polarization and angle, it is possible to predict the region of the scattering hemisphere where the pattern features would be most likely to scatter. Collection optics and a detector placed in this direction could be used to monitor the lines continuously. The specularly reflected light is normally sent to a beam dump.

In patterned wafer inspection it becomes important to differentiate between signal due to particles and defects from the signal due to pattern features on the wafer. Since most particles and defects will occur randomly, it can be assumed that they will not repeat on adjacent die on the wafer, whereas the pattern features will repeat. By subtracting the current die signal from that obtained from an adjacent die, the pattern features can be eliminated. In addition several ADC techniques can be used to segment the defects, cluster them, and identify the defect class to which they belong.

There are several enhancements to the above-mentioned technique that make the patterned wafer defect detection tools sophisticated. First, the PMT signal from each scanned spot on the wafer is amplified and digitized to provide enough dynamic range such that it allows maximum sensitivity for low-scatter regions of the wafer while ensuring that the signal does not saturate in the high-scatter regions. After the entire wafer has been digitized, each region of the wafer with similar scattering characteristics is compared against some threshold value that may be preset during the setup of the tool (50,51). This results in rapid identification of features of interest on the wafer, which can then be subject to the die-to-die comparison for detection of the defects.

Once particles have been detected, it is desirable to obtain a measure of their sizes for subsequent analysis. Inspecting a calibration wafer on the scatter tool accomplishes this. A calibration wafer typically has a uniform deposition of polystyrene latex (PSL) spheres with known diameter. By scanning such a wafer, the size of a detected particle can be estimated from its scattering intensity relative to that of the PSL sphere (52). However, it must be noted that the shape of the particle, its material composition, the film that it lies on, the angle of incidence, and the polarization of the laser beam affect the scattering intensity of the particle. Effectively the particle may look bigger or smaller than its actual size. Nonetheless, it is important to have an estimate because defects below a certain size may not have a detrimental effect on patterned wafers. By neglecting the smaller defects, the task of the yield enhancement engineer is made easier.

AUTOMATIC DEFECT CLASSIFICATION

An in-line microscopy or laser-scattering system provides the user with an electronic "roadmap" of wafer defects. This roadmap is called a "wafermap" and contains a list of all detected defect coordinates and an estimate of defect size. The wafermap is used for subsequent off-line review procedures by providing wafer coordinates to re-locate defects for high-resolution optical, confocal, or SEM analysis. A typical modern fabrication facility, or fab, may have upward of 5000 wafer starts a week, which equates to over 1000 wafermaps that must be analyzed to control the manufacturing process. The amount of data available to yield engineers is quickly outstripping their ability to provide effective and timely yield learning for process characterization and control (36,53).

To address the issue of too much data and too little time, automation technologies in defect detection and review are being enthusiastically developed by universities, federal laboratories, industry consortia, and semiconductor equipment suppliers. This application of artificial intelligence in the wafer analysis community is called automatic defect classification (ADC) on the subdie or defect level and spatial signature analysis (SSA) on the whole-wafer level. These two technologies consider the trees and the forest of wafer analysis, respectively. In this section the concepts of ADC and SSA along with the basic technique of in-line defect detection and off-line re-detection will be described.

Segmentation of Acquired Data

Figure 13 shows the basic steps applied to detecting a wafer anomaly on an in-line inspection tool. A defect is detected on an in-line tool by comparing the wafer die under test with its neighbor. The wafer is scanned in serpentine fashion as shown in Fig. 13(a). Each die under test is compared to its neighboring die through a process of alignment, subtraction, and filtering, as shown in (b) and (c). The defect data are then recorded in the wafermap by defect number (arbitrarily assigned), x and y location, and size prior to database storage shown in (d). Since each die on the wafer contains the same pattern or circuitry, the only difference should be due to local anomalies caused by surface or embedded particle contamination, or extra or missing pattern. These two basic categories of defectivity, namely particle versus pattern, broadly encompass most categories of defects on the wafer, but their description and possible manufacturing sources vary widely. The purpose of automatic detection and classification is to quickly isolate defective areas on the wafer, label and categorize the event, determine the source or cause of the defect, and quickly make a tool or process correction to reduce further yield loss (53).

To facilitate ADC, it is required that the defect be accurately segmented from the wafer pattern. Although the concept of subtractive image processing is easy to comprehend, in practice there are several steps involved. Once the pair of images is obtained, the test image must be registered to the reference image. This requires a shifting of one image to overlay the other to subpixel precision prior to pixel-by-pixel subtraction (54). This is accomplished by selecting one or more registration points that tie one imaged feature to another, and performing a correlation of the two regions to determine small sub-pixel shifts. Normalized correlation is determined by (55),

$$c(x,y) = \frac{\sum_i \sum_j f(x+i, y+j) \cdot g(i,j)}{\sqrt{\sum_i \sum_j f^2(i,j) \cdot \sum_i \sum_j g^2(i,j)}}$$

where $c(x, y)$ is the surface that results from correlating the reference region, $f(x, y)$ with the test region, $g(i, j)$. For example, the intersection of two conductor lines on a device could provide this information. For semiconductor image registration, it is generally required that only a vertical or horizontal translation be determined since the wafer stepper, the device that translates the wafer under the optical system, is constrained to move in these directions. The translation is achieved for the entire image by applying a polynomial interpolation function,

$$x'(x,y) = \boldsymbol{X}^T \boldsymbol{A} \boldsymbol{Y}$$
$$y'(x,y) = \boldsymbol{X}^T \boldsymbol{B} \boldsymbol{Y}$$

where $x'(x, y)$ and $y'(x, y)$ represent the new shifted pixel coordinate pair (x', y') after translation, and the matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ contain the coefficient of the polynomial shift derived from

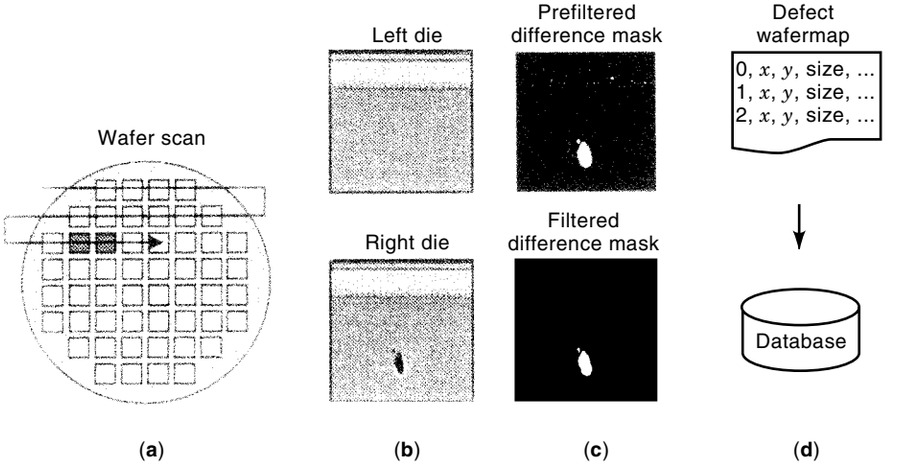

Figure 13. Defect segmentation begins with scanning the wafer row by row (a), collecting neighboring die images (b), comparing and filtering to find anomalies (c), and reporting and data storage (d).

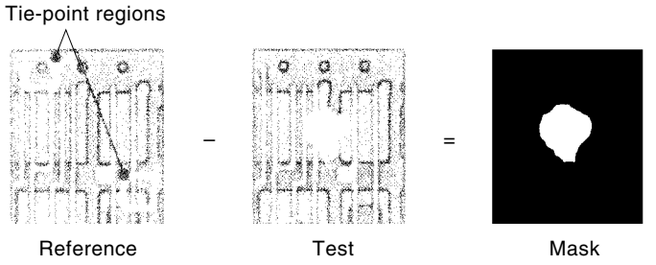

Figure 14. Example of a reference and test die showing the tie-point regions required to measure and correct alignment and the resulting defect mask used for subsequent description.

the correlation points (56). Figure 14 shows an example of a reference and test die image and the resulting defect mask. The mask is a binary image that encompasses the extent of the defect. It is used to extract the centroid location, (x, y), and the defect size (57) prior to inclusion in the wafermap.

Once the alignment has been completed and the images subtracted, a filtering process is applied to the result. The purpose of post filtering is to mitigate residual noise around high contrast areas of the image, such as along conductor or polysilicon lines or along the periphery of transistor regions. The residual noise, which can be considered as "on" pixels in the mask image, can be filtered in a simple manner by rejecting all mask objects below a specified number of pixels in size. This technique is useful and fast but can lead to a discarding of small image anomalies in uniform regions of the die where a defect should be easier to detect. Other filtering techniques that account for device structure can apply stronger filtering near edges and less in open regions, allowing the overall sensitivity of the detection system to be higher. One such technique uses a structural filter derived from the reference image (58). The filter is obtained by digitally differentiating the reference image to give high values in regions of fluctuation and near zero variation in constant regions. This derivative is inverted, normalized, and multiplied by the subtraction result prior to thresholding to generate the defect mask. The result is an attenuation of superfluous data around noisy edges allowing for detection of subtle defects in relatively smooth regions.

Defect Clustering

Defect clustering is a data reduction process. The defectivity on a wafer is recorded as individual events due to the discrete detection process described above. In many instances, very large defects, such as long scratches, large depositions of previous-step material, or chemical stains, are detected as many hundreds, or even thousands, of individual defects. A cluster is a grouping of these individual related defects into a single object. Figure 15 shows several examples of clusters on three different wafermaps. The clusters in (a) are handling scratches; (b) contains a double slot event that is caused by attempting to place a wafer in a slot occupied by another wafer. The wafer in (b) also contains a systematic and distributed pattern that arose from particle contamination in a chemical vapor deposition (CVD) process. The wafer in (c) contains streaks imparted during a spinning process, such as associated with resist deposition or a spin-on-glass process.

The purpose of defect clustering is twofold: (1) to remove clusters from the defect population so that off-line review can focus (primarily albeit not exclusively) on particles, and (2) to remove systematic cluster events so that an accurate count can be made of random particle populations. As a manufacturing process becomes mature, the occurrence of systematic events becomes fewer until ultimately random defects generated by the process tools become the dominant yield-limiting factor. The ability to separate random defects from systematic clusters and distributions is important for predicting future yields from the mature process.

The dominant commercial means of separating clusters from random defects is to form defect groups based on proximity. All defects within a given distance of one another are grouped together and removed from the distributed population as clusters. This technique is useful to a limited extent but results in little or no description of the clusters themselves. For example, a single long scratch event may include many clusters. A group of clusters when pulled together into a larger set is called a spatial signature. A spatial signature is defined as a unique distribution of wafer defects arising from a single manufacturing source (59). SSA is the technology that has been developed to address signature detection and classification. By analyzing distributions of clusters, it is possible to automatically classify and tie the resulting signature directly to a unique manufacturing problem, such as the spin-coater streaks in Fig. 15(c) which are unique to spinning processes. As important as improved cluster analysis is the ability of an SSA approach to classify systematic distributions that do not form clusters, such as the CVD problem shown in Fig. 15(b) which can be recognized and distinguished from a random distribution (60).

Advanced clustering technologies such as SSA are allowing for automatic wafer analysis and data reductions. This automation frees the yield engineer to perform more important work tasks while allowing for 100% wafermap (53). SSA also allows for automated and intelligent subsampling of wafer defects prior to off-line review. For example, it is not generally necessary to review a scratch, since it is known to have been mechanically imparted to the wafer, but a CVD signature may require further analysis to determine the chemical composition of the contaminant. SSA can be applied to filter vari-

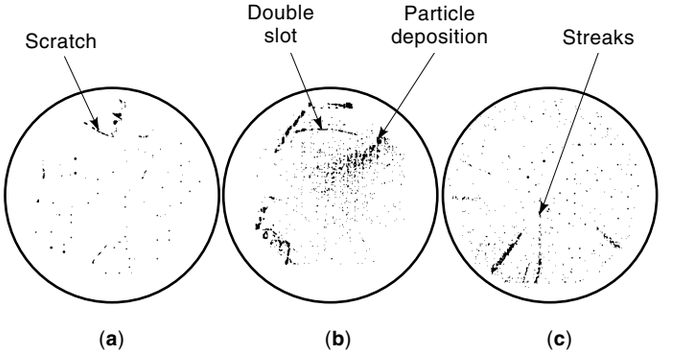

Figure 15. These figures serve as examples of systematic clusters on a series of wafermaps. Each cluster is made up of many individual defects that are correlated to each other based on the manufacturing source.

ous signature types and streamline the off-line review process (61).

Feature-Based Analysis

Both ADC and SSA technologies rely on a feature description of the defect or signature event prior to classification. A feature is a numerical or syntactic description (38) of a segmented object such as a defect or a signature. Common features used to describe semiconductor defects are those based on object shape, or morphology, intensity, color, and texture. Features tend to describe attributes of objects that can be related to human understanding, such as elongation, eccentricity, area, centroid, color components, and texture components. There are also many types of features that are difficult to interpret in human terms but which are very useful for discriminating one object from another. Examples of these include Fourier coefficients (62), wavelet coefficients (63), and high-order object moments, central moments, and invariant moments (64,65). It is possible to completely describe an image object by taking a large number of Fourier, wavelet, or geometric moment coefficients. By taking a small subset of these features, a sufficiently descriptive analysis and subsequent classification can be achieved.

The descriptive features are assembled into a feature vector that is used for subsequent classification (66). A simple example of a two-dimensional feature vector and a feature space is shown in Fig. 16. In the example there are three classes represented that are described by two features, f_x, and f_y. An unknown vector, $\boldsymbol{x}$, is assigned to one of the three classes through the classification process. In the semiconductor environment, a feature vector could contain on the order of one hundred elements rendering viewing of the feature space impossible. This is where the science and art of pattern recognition becomes necessary in the design and test of an effective classifier.

To properly characterize a defect, regardless of whether it is a particle or pattern problem, it is useful to describe the defect in terms of its location on the device structure and what it is overlaying or underlying. For example a particle that causes bridging between two conductors will likely be classified as a "killer" defect, since it impacts electrical function. The same defect on a field area may not electrically impact the device. The classification of these two events would

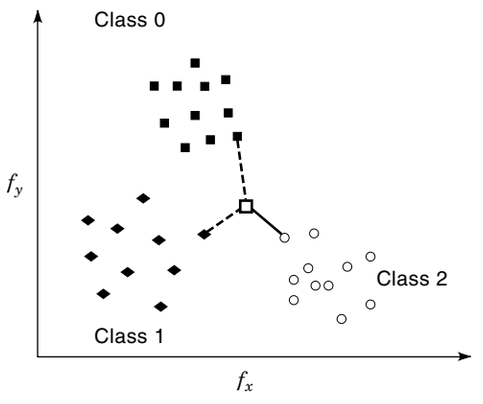

Figure 16. A two-dimensional feature vector, $\boldsymbol{X}$, in a feature space (fx, fy).

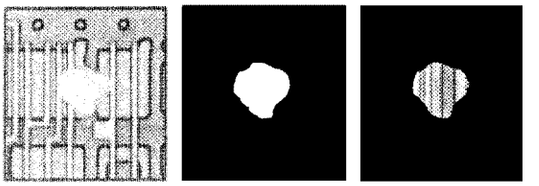

Figure 17. A semiconductor defect (a), defect mask (b), and defect background (c) from the reference image.

therefore require knowledge of the current layer and surrounding material. Feature analysis is extended to consider the layer properties for example, by measuring characteristics of the underlying surface as shown in Fig. 17. The defect in (a) is filtered and a mask describing the defect morphology is found in (b). Properties of the substrate material can be measured by looking in the defect region on the reference image, effectively "under" the defect, as in (c). New approaches to characterizing the underlying device structure are using computer-aided design (CAD) information and advanced image processing to segment the device into, for example, transistor areas, polysilicon lines, conductors, and field areas. Knowing this prior to defect detection can be useful in categorizing the killing potential of a particular defect class. These techniques are computationally intensive at present and are therefore not yet widely used.

Once an adequate description of the defect is obtained whose features are discriminating, such as in the sense of a feature-space representation, a classifier must be designed to automatically assign the object to the appropriate classification category.

Classification

There are several different strategies available to perform automatic classification in the semiconductor-manufacturing environment. For all methods discussed herein, features are extracted from the segmented object, and object background, as described above and are sent to the classifier where a user-defined class label is assigned to the result. The user-defined class label attempts to associate the defect or signature type with a specific manufacturing source, process, or tool.

Classification techniques in general fall into two broad categories: (1) supervised classifiers; and (2) unsupervised classifiers (66). The unsupervised pattern recognition approach assumes no a priori knowledge of the defect classes. The goal of this classifier is to look for patterns in the data by organizing the feature vectors into groups in feature-space based on, for example, a distance metric given by,

$$d(\boldsymbol{x}, \boldsymbol{x}') = \|\boldsymbol{x} - \boldsymbol{x}'\| = [\sum (x - x')^p]^{1/p}$$

where d is the distance between the feature vector $\boldsymbol{x}$ and $\boldsymbol{x}'$ and p is a value such that $p = 1$ is the Manhattan distance, $p = 2$ is the Euclidean distance, and so on (67). For example, the three classes shown in the two-dimensional feature-space of Fig. 16 could easily be grouped using a binary decision tree approach (68). The binary decision tree method will group all vectors according to their pairwise neighbors into a "dendogram" (68). By picking the appropriate cut-point on the den-

dogram, a given number of classes can be selected ranging from one superclass encompassing the entire feature-space, to one class per vector. A cut-point somewhere in between these two trivial cases will provide maximum information. The unsupervised approach does not assume a priori knowledge of the classifications and therefore may or may not map well onto the actual classification groups desired by the user. For this reason unsupervised methods, although good for investigating a given feature-space, are nondeterministic and do not easily map to human-specified classification needs.

Supervised approaches are used exclusively for semiconductor defect classification. Supervised recognition methods rely on the availability of a set of defect or signature examples that are representative of the manufacturing process. These samples are used to train the classifier. The function of the classifier is to partition the feature-space by determining boundaries between the groups of features that correspond to the user-defined classes. These techniques range from statistical approaches such as Bayesian classifiers (69), and neural networks (70), to nonparametric, distance-based methods such as k-means and k-nearest neighbors (k-NN) (71), and finally to fuzzy rule-based approaches (72). There are also combinations of these methods such as the fuzzy k-NN method (73), which has been successfully adapted to the problem of signature classification (74).

For semiconductor pattern recognition, nonparametric classifiers such as the classical *k-NN* and fuzzy rule-based techniques apply well since information about the shape of the distribution of features in the multidimensional space of the classifier is not required. It is difficult to ascertain a statistical parameterization for the large variety of class types encountered. Also in an industrial setting it is often required that the classifier system begins to classify new data with few training examples while providing reasonable accuracy. Bayesian classifiers and neural networks generally require large sample populations to estimate the appropriate statistics and are therefore difficult to implement in general for semiconductor applications. This is primarily due to the diverse nature of the patterns that arise for different manufacturing processes and facilities, coupled with the length of time required to collect large sample populations. Also, over the period of time required to collect large sample sets, acceptable process variations can occur that confuse the boundaries between classes. The nonparametric classifier training set can readily be maintained over time (e.g., by including and excluding examples based on time and date), can be modified often, and can operate with relatively few examples for each class.

Once a feature extraction and classifier approach has been determined, a common method for quantifying the performance and purity of the system is to use a confusion matrix. Table 1 gives an example of a confusion matrix for a simple four-class problem. Note that the confusion matrix can be used to analyze both the training data used to teach the system and the testing data collected by the system in the field.

The left-most column of the table contains the user-defined labels and statistics, e.g., summing all data values across the first row shows that there where 10 examples of class A defects. The second row shows 13 examples of class B defects, and so on. The columns in the matrix represent the classifier results: The first column shows that the classifier placed 10 defect examples in the class A category, the second column shows 15 in the class B category, and so on. The class performance at the end of each row reveals the performance of the classifier in correctly classifying user-defined defect types. The purity metric reveals the percentage of all classified defects that were in the correct class. For example, based on the second row and second column, the classifier correctly classified 92% of class B defects (performance), but of the class B defects classified by the system, only 80% were truly class B (purity). The confusion matrix reveals a great deal of information regarding the strengths and weaknesses of a given classifier system including all of its subcomponents from the extracted features to classifier method to the quality of the training data.

Table 1. Example Confusion Matrix Used to Estimate the Performance and Purity of a Classifier System

	A (classifier)	B	C	D	Class Performance
A (user-defined)	9	1	0	0	90%
B	0	12	0	1	92%
C	0	0	7	0	100%
D	1	2	2	10	67%
Purity	90%	80%	78%	91%	
Total purity 85%			Total performance 87%		

FUTURE CHALLENGES

The ability to quickly identify and eliminate particle sources has resulted in continuous yield improvement for the semiconductor industry, and the yield learning rate improvement on successive process technology nodes has allowed the semiconductor industry to stay on its historical productivity curve (28). However, looking forward, the industry is facing several nonincremental changes (i.e., low/high k dielectrics, copper metallization, dual damascene processing, nonoptical lithography) and continuous improvement of the yield learning rate will become increasingly challenging.

Three factors drive defect inspection technology requirements in the late 1990s. First, the shrinking feature sizes of integrated circuits necessitate sensitivity to smaller defects. In addition, requirements increase throughout as semiconductor processes mature because more area must be scanned in order to detect statistically significant levels of defects. Finally, more rapid yield learning requires shorter learning cycles, where a cycle of learning is defined as the time required to plan an experiment, fabricate the required wafers, test or inspect the wafers, and analyze the data (6,7).

Shrinking Feature Sizes

Integrated circuit feature sizes have been shrinking for more than 30 years, and the 1997 NTRS expects this trend to continue (27,28). Circuits with smaller feature sizes are susceptible to electrical faults induced by smaller defects, requiring defect detection technology to become increasingly more sensitive. Suppliers of defect detection equipment introduce new technology to keep up with these requirements. For example, suppliers of optically based defect inspection tools introduce new light sources with shorter wavelengths and optimize the lens optics for the new physical conditions. This trend is likely

to continue until inspection equipment needs to capture defects so small that optically based detection is no longer suitable for the job.

As IC feature sizes shrink, semiconductor processes become more complex, and new defect classes become yield limiters. Figure 18, for example, depicts the systematic failure modes of copper dual damascene, the anticipated choice of global interconnect technology for the 180 nanometer node. New process technologies like CMP introduce hitherto unknown classes of surface defects (25), whereas other failure modes reside up to 1500 nm below the surface where they are exceedingly difficult to detect. Metal voids and barrier metal punch through may even elude optical inspection equipment completely.

The increased number of defect types is driving up the cost of fault reduction for semiconductor manufacturers. Yield engineers have observed that different types of defect inspection equipment capture some defect types more effectively than they do others (75,76). Companies are thus forced to purchase multiple defect inspection tool types, in order to identify most potential sources of electrical faults. As process complexity increases, defect detection may pose a daunting economic challenge to all but the largest semiconductor manufacturers.

Process Life Cycle

The performance requirements for defect detection on patterned wafers depend upon the maturity of the semiconductor process. Figure 19 shows that the sensitivity requirement of patterned wafer inspection decreases when a semiconductor process moves from research and development, through the yield ramp, to volume production. The throughput requirement increases by more than two orders of magnitude from the beginning of PR&D phase to the beginning of VP phase.

The combined requirements of shrinking feature sizes and process life cycle segment the market for inspection tools. Shrinking feature sizes drive sensitivity, or the ability to see small defects, while throughput requirements increase as an IC process matures. Inspection tools today typically focus on either sensitivity or throughput, with very little latitude for in-line (i.e., in a single tool) trade-off of those performance parameters. This means different tools (or different recipes on those tools) are used for different phases of manufacturing process. IC manufacturers must purchase multiple inspection tool types, which drives up their cost of process development. Clearly more agile defect inspection tools would help the semiconductor industry control cost.

Shorter Cycles of Learning

Capital productivity drives the trend toward shorter data cycles. IC manufacturers invest billions of dollars in process equipment, and they are interested in obtaining as rapid a return on their investment as possible. Rapid yield learning is thus becoming an increasingly important source of competitive advantage in the semiconductor industry. The sooner potentially lucrative circuits yield, the sooner the manufacturer can generate a revenue stream. Conversely, rapid identification of the cause of yield loss can restore a revenue stream and prevent the destruction of material in process (77).

Historically IC manufacturers have obtained their electrical fault data from microelectronic test structures. However, today they increasingly rely on in-line data from patterned wafer inspection, which can detect yield-limiting defects on shorter cycles. In-line data provide a cumulative count of defects throughout the process. Subtracting the previously detected defects from the current image identifies defects that have occurred since the last inspection. Thus the source of recently detected defects can be localized more effectively.

Recent studies have shown that shortening the defect learning cycles accelerates yield learning by increasing experimentation capacity. Defects must be detected, analyzed, and eliminated within increasingly shorter time periods. Consequently successful yield improvement tends to consist of a total systems approach that involves electrical testing, defect inspection and in situ fault detection. A defect reduction team can thus develop true yield management capability by correlating data obtained from methods with short data cycles to those extracted from methods with longer ones. Once defect databases become large enough, signals from short-cycle methods can foreshadow effects on final yield (4–6,8,78,79).

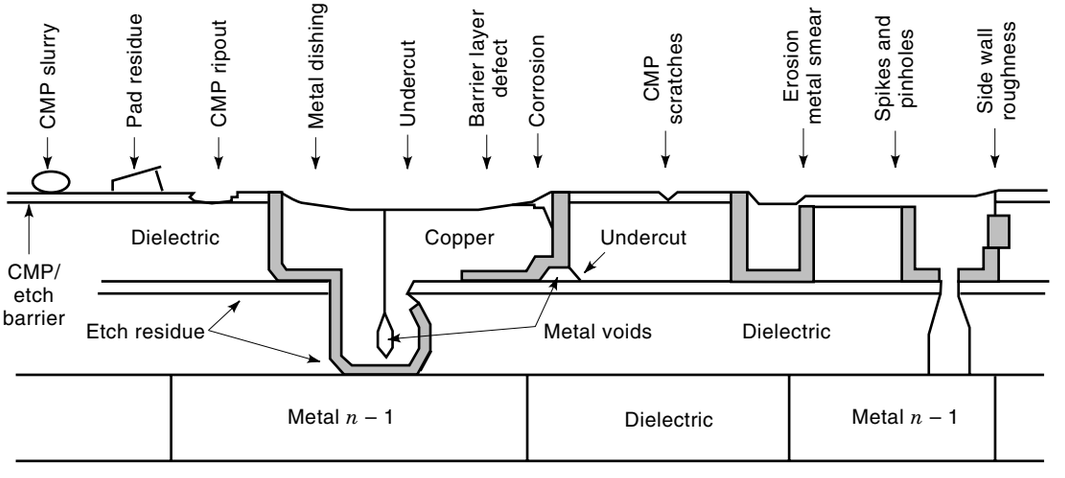

Figure 18. A cross section of most systematic failure modes that may cause defects in dual damascene processing of wafers.

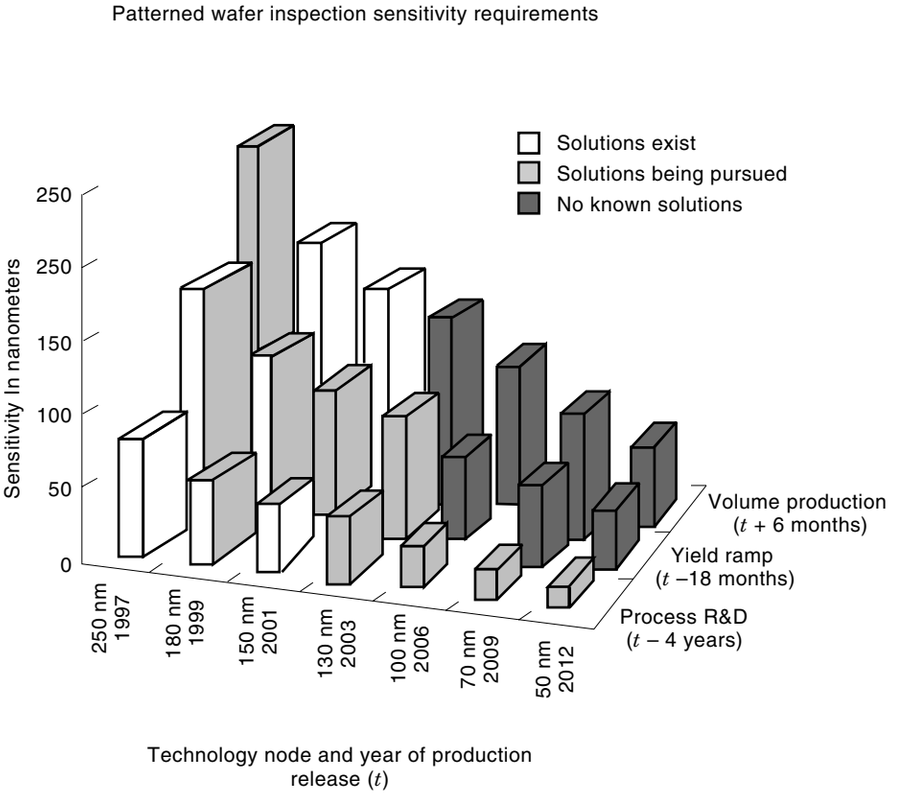

Figure 19. NTRS-1997 sensitivity requirements for patterned wafer inspection.

Figure 20 compares the relative data cycles of key inspections of copper dual damascene, the likely choice for global interconnect for the 180 nm technology node. The horizontal axis in Fig. 20 lists the process steps of copper dual damascene, which may repeat between 5 to 10 times within a full VLSI process, making global interconnect the process module with the highest number of defects. However, similar patterned wafer inspections occur in other modules of a VLSI process.

High-aspect ratio inspection (HARI) is a special challenge, because defects that reside below the wafer surface are harder to capture. HARI sensitivity requirements are thus tougher than other forms of patterned wafer inspection. The HARI challenge occurs at many stages of the fabrication cy-

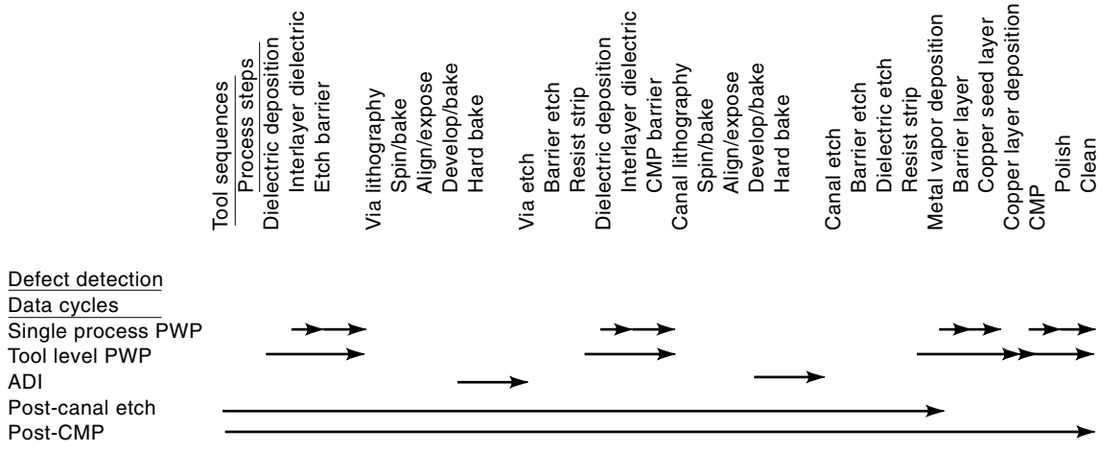

Figure 20. Defect detection data cycles of copper dual-damascene based on the SEMATECH process flow for the 180 nm technology node. PWP, ADI, and CMP respectively stand for particles per wafer per pass, after-develop inspection, and chemical mechanical polish. Indented process steps can occur within one of many chambers of a process tool or within a process tool that is part of a tool cluster. A sequence of process steps that occurs within one cluster tool or a cluster of process tools constitutes a tool sequence.

cle, including shallow trench isolation and local interconnect. The first-level contacts exhibit the highest aspect ratios encountered during any inspection. They are likely to exceed 6 by 1999 and grow to 12 by 2012. Adjacent first-level contacts may also vary in depth because they may contact with both polysilicon and the active area, which are not coplanar. Dual damascene structures, which are likely to prevail by the year 2000, consist of combinations of canals that contain interconnect lines and vias between interconnect layers. Their combined aspect ratios may exceed five by 2006 (26).

The difficulty of detecting defects in vias is exacerbated by the fact that there will be typically 10^{11} vias present on a 300 mm product wafer per layer of vias. Defects in 10 to 20 of those 10^{11} vias would have a significant detrimental effect on yield. Directing the interrogating agents (photons, electron, ions, etc.) onto the bottom of the narrow vias adds to the difficulty of detecting such defects, and having the signal energy make its way back out of the via to a detection system compounds it. There are no methods for which HARI feasibility has been conclusively demonstrated, even at the Process Research and Development phase, and certainly not at the Volume Production phase.

The data cycles of different defect inspections vary greatly. For example, an inspection at the head of an arrow in Fig. 20 evaluates all the process steps between the head and the tail of the arrow. Thus the post-CMP inspection covers all process steps of dual damascene. Many wafer fabs also intend to conduct a post-canal-etch inspection, an example of HARI that localizes the sources of defects to fewer process steps. In absence of such short-cycle inspections, a via etch problem could go unnoticed for more than a day, which would put a large number of wafers at risk in a volume production facility. After-develop inspection may even capture systematic lithography defects in time to perform rework, which can increase line yield significantly and prevent yield crashes through early defect detection.

The final major challenge of shortening defect data cycles consists of providing real-time defect/fault classification capabilities for in situ monitors and defect inspection equipment. At the present time, classifying defects detected by optical means or by SEM is relegated to defect review stations, which have to re-detect the defect found by the inspection tool. Thus defect characterization becomes a two-step process consisting of defect detection and defect review. In addition optical defect review is becoming increasingly difficult because numerous killer defects are no longer detectable by optical means. According to the NTRS, even the 250 nm technology requires SEM-based defect review and classification solutions.

Table 2 outlines the technology requirements for automatic defect classification (ADC). For some yield monitoring applications (mostly back end metal layers) in high-volume manufacturing, optical based redetection and ADC will be extendible into the 150 nm and 130 nm technology nodes. However, for yield ramp engineering applications (30% to 70% minimum feature size resolution) at these nodes and front end of line applications at the 250 nm and 180 nm nodes, SEM-based solutions are necessary as optical redetection and ADC falls off dramatically at ~250 nm. In this regard SEM-based review and ADC solutions will be necessary in many applications (at earlier technology nodes for front end of line and later for back end of line). Detection tools which provide in-line ADC capability based either on image or light scatter analysis will also be necessary to assist in accelerating cycles of learning, as well as optimizing SEM-based tool utilization. It is also believed that a combination of optical and SEM-based solutions will be necessary to classify those defects not redetectable via SEM (e.g., previous layer defects) and those not redetectable via optics ($\leq$ 250 nm surface/pattern defects).

In addition to the application challenges described above, inspection technologies themselves face several hurdles in future development. Numerical aperture of the lenses and

Table 2. Technology Requirements for Automatic Defect Classification and Defect Review as Outlined in NTRS-1997

Year of First Product Shipment Technology Generation	1997 250 nm	1999 180 nm	2001 150 nm	2003 130 nm	2006 100 nm	2009 70 nm	2012 50 nm
Automatic Defect Classification							
Resolution (nm)	125	90	75	65	50	35	25
Detectability (% redetection)	95	97	98	99	99	99	99
Accuracy (% of expert classification)	85	90	90	95	95	95	95
Repeatability (%)	95	97	98	99	99	99	99
Reproducibility (cOV%)	5	4	4	3	2	2	2
Speed—optical (s/defect)	3	2	2	1	1	1	1
Speed—SEM (s/defect)	20	10	10	5	5	5	5
Speed—SEM w/elemental (s/defect)	35	25	20	15	10	10	10
Defect Review							
Resolution (nm)	125	90	75	65	50	35	25
Coordinated accuracy @ max sensitivity (μm)	3	3	2	2	1	1	1

Note: Solutions exist for the specifications that are in unshaded cells, while gray shading indicates that solutions are being pursued. There are no known solutions to the specifications listed with white letters on black background.

achievable pixel sizes of CCD cameras limit imaging tools to a maximum resolution of about 150 nm. Laser-scattering tools have taken advantage of the wavelength scaling principle, whereby sensitivity to the intensity of scattering from a smaller particle can be obtained by illuminating the particle with shorter wavelength light. In the future, as contaminant sizes approach closer to the dimensions of the inherent surface roughness of the patterns and films on the wafer, the detectibility of such particles will become questionable. Improvements in laser technology should target the development of compact and reliable deep-ultraviolet sources. New detection schemes that take advantage of optical properties other than intensity of scatter will have to be developed in parallel.

Moreover, improving the resolution increases the demand on the data rate of the inspection system. Significant breakthroughs in electronics have afforded rates of 400 million pixels per second in present systems; however, the limits of quantum physics are being stretched to acquire signals as rapidly as they can be imaged onto the pixels of TDI cameras. Enhanced CCD and PMT sensors with higher photosensitivity will be required soon. Future inspection systems could include subsystems with parallel imaging of different parts of a wafer using multiple cameras and/or detectors, each linked to image analysis engines. Advanced computing architecture that facilitates rapid acquisition, intelligent analysis, and compact storage of inspection data will be an essential component of such systems. All these issues will have to be tackled with marginal increments in tool costs to keep in tune with lowering the overall cost of producing better-quality semiconductor devices.

CONCLUSIONS

While investigating defect inspection technology in the semiconductor industry, the authors of this article came to the conclusion that chip makers and their tool suppliers exhibit asymmetric innovation patterns. The chip makers innovate by yield learning, in a manner consistent with the economics of experimentation (7) and the economics of products with low unit cost. They also tend to repeat this pattern for every process technology node. The tool suppliers, in a manner consistent with the economics of developing products with large unit costs, improve their products by reconfiguring them (80). They replace subsystems to improve performance, and they rely on leading-edge chip makers to use a tool in order to improve its quality (81). This typically involves a problem-solving process that iterates between the chip makers and the tools suppliers.

The chip maker specifies the parameter of the tool based on need, and the tool supplier develops a prototype to the chip maker's specifications. The user evaluates the prototype and specifies changes and improvements as evidence dictates. This process continues until the user is satisfied. The ease of use and ease of transfer of the information output of equipment is critical process because it enhances the chip maker's innovation capability (20).

The creation of defect inspection tools presents a special challenge for the previously described asymmetric innovation pattern. Chip makers, when they start working on a new technology node, will specify a tool for the process R&D environment, stressing increased sensitivity and resolution as the primary requirements. The tool suppliers will most likely respond by reconfiguring the most sensitive tool that they have on the market. The reconfigured tool will most likely provide a nonoptical solution that meets the chip maker's demands for sensitivity and resolution. Throughput in the early stages of process R&D is not a major consideration.

However, as time passes, the chip maker will improve the quality of its process and demand a tool with greater throughput. Since the throughput requirements for defect inspection increase exponentially with the maturity of the process, the tool suppliers will not be able to increase the scan speed at a rate that will satisfy the chip makers. The tool suppliers will instead respond by reconfiguring an existing tool, based on a completely different technology such as optical imaging. This new tool will meet the new throughput requirements and provide more sensitivity than any previous imaging tool, but it will most likely not be able to match the R&D tool's sensitivity. Once the semiconductor process reaches volume production, the chip maker will demand another quantum leap in throughput, and the suppliers will once again respond with reconfiguring an existing tool based on yet another technology, possibly laser scattering. Once again, the laser-scattering tool, while meeting the throughput requirements and exceeding the sensitivity requirements of the previous technology node may provide less sensitivity than the imaging tool it replaces. In summary, it is much easier for a tool supplier to tweak the sensitivity of a tool by reconfiguring its architecture, than it is to increase the tool's throughput. No tool supplier is therefore likely to provide an inspection tool that follows a semiconductor process through all phases of development, at least not in a configuration that any chip maker can afford.

The expected high price tag of defect inspection tools may cause many companies to reevaluate the cost/benefit relationships of defect detection. For instance, most of today's cost of ownership models do not include defect inspection. Consequently semiconductor manufacturers are continually tempted to underinvest in inspection tools. They argue that inspection adds no direct value to the wafers and focus their resources elsewhere. However, a suboptimal number of inspection steps introduces substantial risk of yield loss, whose cost has to be balanced against the cost of ownership of the inspection equipment (82). Inspection tools, no matter how expensive, are therefore here to stay.

ACKNOWLEDGMENTS

The authors would like to thank David Jensen for reviewing this article.

BIBLIOGRAPHY

1. K. J. Arrow, The economic implications of learning by doing, *Rev. Econ. Studies,* **29**: 155–173, 1962.
2. T. P. Wright, Factors affecting the cost of airplanes, *J. Aeronaut. Sci.,* **3**: 122–128, 1936.
3. J. M. Juran, ed., *Quality Control Handbook,* 3rd ed., New York: McGraw-Hill, 1974.

4. C. Stapper and R. Rosner, Integrated circuit yield management and yield analysis: Development and implementation, *IEEE Trans. Semicond. Manuf.,* **8**: 95–101, 1995.
5. C. Weber, B. Moslehi, and M. Dutta, An integrated framework for yield management and defect/fault reduction, *IEEE Trans. Semicond. Manuf.,* **8**: 110–120, 1995.
6. C. Weber, *Accelerating three-dimensional experience curves in integrated circuit process development,* Management of Technology master's thesis, Massachusetts Institute Technology, 1996.
7. S. H. Thomke, Managing experimentation in the design of new products, *Harvard Business School Working Paper* 96-037. *Management Science,* in press.
8. D. Jensen, C. Gross, and D. Mehta, New industry document explores defect reduction technology challenges, *Micro* **16** (1): 35–44, 1998.
9. D. L. Dance, D. Jensen, and R. Collica, Developing yield modeling and defect budgeting for 0.25 um and beyond, *Micro* **16** (3): 51–61, 1998.
10. P. Cielo, *Optical Techniques for Industrial Inspection,* New York: Academic Press, 1988.
11. B. G. Batchelor, D. A. Hill, and D. C. Hodgson, *Automated Visual Inspection,* Amsterdam, The Netherlands: North-Holland, 1985.
12. L. F. Pau, *Computer Vision for Electronics Manufacturing,* New York: Plenum, 1990.
13. J. E. Thompson, *Inspection Organization and Methods,* New York: McGraw-Hill, 1950.
14. E. S. Roth and R. R. Runck, *Functional Inspection Techniques,* Dearborn, MI: American Soc. of Tool Manuf. Engineers, 1967.
15. C. W. Kennedy and D. E. Andrews, *Inspection and Gaging,* 4th ed., New York: Industrial Press, 1967.
16. R. T. Chin and C. A. Harlow, Automated visual inspection: A survey, *IEEE Trans. Pattern Anal. Mach. Intell.,* **4**: 557–573, 1982.
17. R. T. Chin, Automated visual inspection: 1981 to 1987, *Comput. Vis., Graph. Image Process.,* **41**: 346–381, 1988.
18. E. Bayro-Corrochano, Review of automated visual inspection 1983 to 1993, Part I: Conventional approaches, *Proc. SPIE,* **2055**, *Intell. Robots Comput. Vision XII,* 1993, pp. 128–158.
19. E. Bayro-Corrochano, Review of automated visual inspection 1983 to 1993, Part II: Approaches to intelligent systems, *Proc. SPIE,* **2055**, *Intell. Robots Comput. Vision XII,* 1993, pp. 159–172.
20. E. von Hippel, "Sticky information" and the locus of problem solving: Implications for innovation, *Manag. Sci.,* **40**: 429–439.
21. D. Reinertsen, Who dunit? The search for new-product killers, *Electron. Bus.* (July): 62–66, 1983.
22. P. Smith and D. Reinertsen, *Developing Products in Half the Time,* New York: Van Nostrand Reinhold, 1991.
23. S. Weelwright and K. Clark, *Revolutionizing Product Development,* New York: Free Press, 1992.
24. G. Moore, Progress in digital integrated circuits, *IEDM Tech. Dig.,* 1975, p. 11.
25. C. Dennison, Developing effective inspection systems and strategies for monitoring CMP processes, *Micro* (February): 31–41, 1998.
26. Semiconductor Industry Association, *The National Technology Roadmap for Semiconductors,* 1997, pp. 163–178.
27. Semiconductor Industry Association, *The National Technology Roadmap for Semiconductors,* 1997, pp. 99–113.
28. C. Weber, D. Jensen, and E. D. Hirleman, What drives defect detection technology? *Micro* (June): 51–72, 1998.
29. U. Kaempf, Statistical significance of defect density estimates, *Proc. IEEE/ICMTS,* 1988, pp. 107–113.
30. W. Tang, Circuit integation pushes image sensor performance, *Laser Focus World,* **33** (2): 129–139, 1997.
31. S. Kempainen, CMOS Image Sensors: Eclipsing CCDs in visual information? *EDN,* **42** (21): 105–119, 1997.
32. E. Fossum, Image capture circuits in CMOS, *Int. Symp. VLSI Technol. Syst. Appl.,* 1997.
33. *CCD Image Sensors Databook,* DALSA, Ontario, Canada, 1996.
34. T. S. Newman and A. K. Jain, A survey of automated visual inspection, *Comput. Vis. Image Understand.,* **61** (2): 231–262, 1995.
35. S. G. Chamberlain, High speed, low noise, fine resolution, TDI CCD imagers, *SPIE CCDs Solid State Opt. Sensors,* 1990.
36. *The National Technology Roadmap for Semiconductors: Technology Needs,* San Jose, CA: Semiconductor Industry Association, 1997, p. 163.
37. L. Reimer, *Image Formation in Low-Voltage Scanning Electron Microscopy,* vol. 12, Bellingham, WA: SPIE Optical Engineering Press, 1993.
38. E. I. Cole Jr. and J. M. Soden, Scanning electron microscopy techniques for IC failure analysis, *Microelectronic Failure Analysis,* 3rd ed., Materials Park, OH: ASM International, p. 163.
39. R. C. Gonzalez and P. Wintz, *Digital Image Processing,* 2nd ed., Reading, MA: Addison-Wesley, 1987, p. 94.
40. J. D. Gaskill, *Linear Systems, Fourier Transforms, and Optics,* New York: Wiley, 1978, p. 269.
41. Lord Rayliegh, *Proc. Royal Soc.,* **A79**: 399, 1907.
42. C. E. Bohren and D. R. Huffman, *Absorption and Scattering of Light by Small Particles,* New York: Wiley, 1983.
43. T. A. Germer and C. C. Asmail, Bidirectional ellipsometry and its application to the characterization of surfaces, *Proc. SPIE,* **3121**: 1997, pp. 173–182.
44. T. A. Germer, Angular dependence and polarization of out-of-plane optical scattering from particulate contamination, subsurface defects, and surface microroughness, *Appl. Opt.,* **36** (33): 8798–8805, 1997.
45. A. J. Hunt and D. R. Huffman, A new polarization-modulated light scattering instrument, *Rev. Sci. Instrum.,* **44** (12): 1753, 1973.
46. J. C. Stover, *Optical Scattering: Measurement and Analysis,* 2nd ed., Washington: SPIE Press, 1995.
47. C. T. Larson, Measuring haze on deposited metals with light-scattering-based inspection systems, *Micro* **14** (8): 31–38, 1996.
48. J. A. Bender, T. D. Henning, and M. L. Bernt, Near specular measurements of integrated scatter, *Proc. SPIE,* **1753**: 1992, pp. 121–126.
49. R. S. Howland, K. B. Wells, and B. M. Trafas, High-speed detection of pattern defects using laser scattering, *Solid State Technol.,* **38** (11): 123–126, 1995.
50. B. M. Trafas et al., Extendibility of laser scanning tools for advanced wafer inspection, *Proc. SPIE,* **2439**: 1995, pp. 164–173.
51. K. A. Klicker et al., Practical reduction of instrument signature in near specular light scatter measurements, *Proc. SPIE,* **818**: 1987, pp. 26–33.
52. R. S. Howland et al., Using calibration curve modeling for scanning surface inspection systems, *Micro* **13** (7): 61–71, 1995.
53. K. W. Tobin et al., Automated analysis for rapid defect sourcing and yield learning, *Future Fab Int.,* **1**: London: Technology Publishing, 1997, p. 313.
54. K. W. Tobin, S. W. Allison, and J. K. Jordan, Automated pattern recognition: Video weld authentication system, *Applications of Digital Image Procesing XII, Proc. SPIE,* **1153**: 1989, p. 412.
55. J. C. Russ, *The Image Processing Handbook,* Boca Raton, FL: CRC Press, 1992, p. 219.
56. S. S. Gleason, M. A. Hunt, and W. B. Jatko, Subpixel measurement of image features based on paraboloid surface fit, *Proc. SPIE,* **1386**: 1990.

57. A. K. Jain, *Fundamentals of Digital Image Processing,* Englewood Cliffs, NJ: Prentice-Hall, 1989, p. 377.
58. W. B. Jatko, M. A. Hunt, and K. W. Tobin, Nonlinear filter derived from topological image features, *SPIE Aerospace Conf.,* Orlando, FL: 1990.
59. K. W. Tobin et al., An image paradigm for semiconductor defect data reduction, *SPIE's 21st Annu. Int. Symp. Microlithography,* Santa Clara, CA, 1996.
60. K. W. Tobin et al., Automatic classification of spatial signatures on semiconductor wafermaps, *SPIE 22nd Annu. Int. Symp. Microlithography,* Santa Clara, CA, 1997.
61. S. S. Gleason, K. W. Tobin, and T. P. Karnowski, An integrated spatial signature Analysis and automtic defect classification system, *191st Meet. Electrochem. Soc.,* 1997.
62. M. O. Freeman and B. E. A. Saleh, Moment invariants in the space and frequency domains, *J. Opt. Soc. Amer. A,* **5** (7): 1073, 1988.
63. L. Prasad and S. S. Iyengar, *Wavelet Analysis with Applications to Image Processing,* Boca Raton, FL: CRC Press, 1997.
64. C. Teh and R. T. Chin, On image analysis by methods of moments, *IEEE Trans. Pattern Anal. Mach. Intell.,* **10** (4): 496, 1988.
65. S. X. Lio and M. Pawlak, On image analysis by moments, *IEEE Trans. Pattern Anal. Mach. Intell.,* **18** (3): 252, 1996.
66. J. T. Tou and R. C. Gonzalez, *Pattern Recognition Principles,* Reading, MA: Addison-Wesley, 1974.
67. R. O. Duda and P. E. Hart, *Pattern Recognition and Scene Analysis,* New York: Wiley, 1973.
68. C. H. Chen, L. F. Pau, and P. S. P. Wang, *Handbook of Pattern Recognition and Computer Vision,* River Edge, NJ: World Scientific, 1993, p. 16.
69. J. M. Bernardo and A. F. M. Smith, *Bayesian Theory,* New York: Wiley, 1994.
70. S. Haykin, *Neural Networks: A Comprehensive Foundation,* New York: Macmillan, 1994.
71. K. Fukunaga, *Introduction to Statistical Pattern Recognition,* 2nd ed., New York: Academic Press, 1990.
72. Y. Yuan and M. J. Shaw, Induction of fuzzy decision trees, *Artificial Intell. Tech. Rep.,* UIUC-BI-AI-DSS-94-02, Univ. Illinois, Urbana-Champaign, 1994.
73. J. M. Keller, M. R. Gray, and J. A. Givens, A fuzzy k-nearest neighbor algorithm, *IEEE Trans. Syst., Man, Cybern.,* **15** (4): 580, 1985.
74. K. W. Tobin, S. S. Gleason, and T. P. Karnowski, Adaptation of the fuzzy k-nearest neighbor classifier for manufacturing automation, *SPIE's 10th Int. Symp. Electron. Imaging: Sci. Technol.,* San Jose, CA, 1998.
75. A. Shapiro, T. James, and B. Trafas, Advanced inspection for 0.25-μm-generation semiconductor manufacturing, *SPIE's 22nd Annu. Int. Symp. Microlithography,* Santa Clara, CA, 1997, pp. 445–451.
76. M. Seliger and M. B. Nasr, Inspecting wafers using the Orbot-W720 at metal ASI, *Process Diagnostic and Control User Group Meet.,* Santa Clara, CA, 1998.
77. P. Silverman, Capital productivity: Major challenge for the semiconductor industry, *Solid State Technol.,* **37** (3): 104, 1994.
78. M. Iansiti and J. West, Technology integration, *Harvard Bus. Rev.* May–June: 69–79, 1997.
79. C. Weber, J. Rosner, and P. Chang, The role of self-directed work teams in a concurrent, multi-generation process development effort, *Proc. IEEE/SEMI/ASMC,* Cambridge, MA: 1997, pp. 322–327.
80. R. Henderson and K. Clark, Architectural innovation: The reconfiguration of existing product technologies and the failure of established irms, *Admin. Sci. Quart.,* **35**: 9–30.
81. E. von Hippel, *The Sources of Innovation,* Oxford, UK: Oxford University Press, 1988.
82. E. Wang et al., Valuation of yield management investments, *Proc. IEEE/SEMI/ASMC,* Cambridge, MA, 1997, pp. 1–7.

VIJAY SANKARAN
CHARLES M. WEBER
FRED LAKHANI
SEMATECH
KENNETH W. TOBIN, JR.
Oak Ridge National Laboratory

INSTALLATION OF COMPUTERS. See COMPUTER INSTALLATION.

INSTANTANEOUS FREQUENCY ESTIMATION. See FREQUENCY MODULATION.

INSTRUCTION, COMPUTER AIDED. See COMPUTER-AIDED INSTRUCTION.