Model-Based Shape From Shading for Microelectronics Applications

A. Nissenboim,[1] A. M. Bruckstein[2]

[1] Applied Mathematics Department, Technion–Israel Institute of Technology, Technion City, Haifa 32000, Israel

[2] Computer Science Department, Technion–Israel Institute of Technology, Technion City, Haifa 32000, Israel

Received 20 June 2006; accepted 10 August 2006

ABSTRACT: Model-based shape from shading (SFS) is a promising paradigm introduced by Atick et al. [Neural Comput 8 (1996), 1321–1340] in 1996 for solving inverse problems when we happen to have a lot of prior information on the depth profiles to be recovered. In the present work we adopt this approach to address the problem of recovering wafer profiles from images taken using a scanning electron microscope (SEM). This problem arises naturally in the microelectronics inspection industry. A low-dimensional model, based on our prior knowledge on the types of depth profiles of wafer surfaces, has been developed, and based on it the SFS problem becomes an optimal parameter estimation. Wavelet techniques were then employed to calculate a good initial guess to be used in a minimization process that yields the desired profile parametrization. A Levenberg–Marguardt (LM) optimization procedure has been adopted to address ill-posedness of the SFS problem and to ensure stable numerical convergence. The proposed algorithm has been tested on synthetic images, using both Lambertian and SEM imaging models. © 2006 Wiley Periodicals, Inc. Int J Imaging Syst Technol, 16, 65–76, 2006; Published online in Wiley InterScience (www.interscience.wiley.com). DOI 10.1002/ima.20065

Key words: shape from shading; model-based SFS; wafer inspection; SEM images

I. INTRODUCTION

The problem of recovering a 3D object's shape from its shaded image, namely the shape from shading (SFS) problem, has intrigued the computer vision researchers for almost 30 years. The research in this area was mostly inspired by the fact that our brain has an outstanding ability to perceive the depth of the observed scene from 2D images on the retina. Shading is only one of the clues used by our brain to do the job. However, in words of Marr (1982), "... one of the most interesting problems in the theory of human early vision, along with color, is exactly what and how much information we are able to recover from shading." It is interesting that despite the lack of our understanding of this cognitive feature and major difficulties of our mathematical modeling attempts, nowadays there is a need for SFS-based applications in some branches of the computer industry. Manufacturing of integrated circuit wafers is an expensive and delicate process and a great deal of effort is invested in the control of quality and acceptability of such crucial device features as contact holes, tracks, etc. Typically, the manufacturers are interested in measuring the geometrical parameters of the features and due to the "nanosizes" of the structures studied, the nondestructive low-voltage scanning electron microscope (SEM) is used. Thus, the necessity of 3D surface measurements from 2D SEM images naturally arises. SEM inspection machines have different configurations. If an image has been taken by a SEM machine with single electron detector, then one faces the issues of monocular SFS. In case of 2 or 4 detectors the shape from photometric stereo-algorithms are the natural choice. The very nature of the SFS idea is to exploit the fact that the variations of surface's orientation cause the variations of the brightness (which we call "shading") at corresponding areas in the image. However, the brightness only carries information about the projection of surface normal on the light source direction and hence the surface normal at each point cannot be uniquely discovered. In fact, infinitely many normal vector fields could, in principle, be associated with any given image. Thus, mathematically speaking, the SFS problem appears to be ill-posed, and since the mathematical formulation of the problem turns out to be a nonlinear first-order PDE, neither existence nor uniqueness of the solution is ensured. One way to overcome these fundamental difficulties is to reformulate the SFS as a variational problem while introducing additional smoothness and integrability, or regularization terms into the minimized functional. Unfortunately, this approach has some serious drawbacks. In this work, following a paradigm introduced by Atick et al. (1996), we try to directly use a priori information about the geometry of the studied surfaces. Giving up generality, we assume a model of the surface controlled by a finite set of parameters. The resulting parameter estimation SFS problem appears to be simpler than that of the original one. Simple

Correspondence to: A. Nissenboim; E-mail: alexn@il.ibm.com

analytical observations reveal that good initial estimates are vital for our method. The initial estimates are obtained by means of analysis of the wavelet decomposition of the given image. Moreover, our method benefits from another important simplification, due to the wavelet decomposition of the image data. It enables us to localize image singularities, which correspond to surface features. The parameter fitting process can be carried out independently and in parallel for every feature. Finally, to address the ill-posedness issue, an iterative Levenberg–Marquardt (LM) minimization procedure has been adopted in our work. This paper is organized as follows. In Section II, we mathematically formulate the SFS problem. In Section III, we briefly overview the previous relevant work. Section IV contains the description of the new method. Section V presents results and discussion.

II. THE SFS PROBLEM

A. General Problem Setup. The monocular SFS problem is defined as follows. Suppose we are given a smooth height field $z = z(x,y)$ over some region $D = \{(x,y) \mid (x,y) \in R^2\}$. Naturally, its shaded image displays variations of brightness and is represented as the intensity map $I(x,y)$. The value of $I(x,y)$ at each point (x,y) depends on reflectance properties of the surface, its gradient, and imaging geometry parameters like light direction, and so on. This dependence is often called reflectance function, and we denote it as $R = R(p,q)$, where $p = z_x$ and $q = z_y$. The relationship

$$I(x,y) = R(p(x,y), q(x,y)) \tag{1}$$

is called the irradiance equation, and we state the SFS problem as an attempt to recover the surface height field $z(x,y)$ from a single shaded image $I(x,y)$ given the reflectance function R, i.e., to determine $z(x,y)$ that satisfies the irradiance equation (1).

Some assumptions are to be made about imaging geometry. First, we assume that the size of the studied object is small, compared to the viewing distance, which enables us to presume orthographic projection to the image plane. We also assume that the camera direction coincides with the Z-axis. In this case one can choose the coordinate system of both image and object planes to be identical and denoted by (X,Y). We denote by η the composite albedo, which consists of intensity of the light source and the ratio of the total reflected light flux to the total incident light flux on the object's surface. We assume that η is constant along the surface. We also denote $\vec{L}$ to be the unit vector of the illumination direction. We assume that there is only one source of light located at infinity.

B. Imaging Models. In this work we used two models of image formation—a Lambertian model and a simple model of SEM reflectance. We say that a surface exhibits Lambertian diffuse reflection property if under uniform illumination the value of $I(x,y)$ is proportional to the cosine of the angle between the normal to the surface at the point (x,y) and the light direction. Thus,

$$R(p,q) = \eta \langle \vec{n}(x,y), \vec{L} \rangle$$

where $<,>$ is a standard inner product and

$$\vec{n}(x,y) = \frac{(-p(x,y), -q(x,y), 1)}{\sqrt{p^2(x,y) + q^2(x,y) + 1}}$$

is a unit normal vector to the surface $z = z(x,y)$. In case of SEM, the simplest image formation model is given by

$$R(p,q) = \frac{\eta}{\langle \vec{n}(x,y), \vec{L} \rangle}$$

This model holds well when specimen is coated with gold and in absence of charging artifacts. For rigorous theoretical foundations of SEM image formation see Reimer (1993). It turns out that the problem of computing η and $\vec{L}$ can be solved separately and, therefore, we assume them to be known (see Zheng and Chellappa, 1991). This allows us to concentrate our efforts on solving the SFS problem itself. We should note here that the algorithm presented below in no way depends on the particular imaging model.

III. PREVIOUS WORK

A. Horn's Variational Approach. In this section we would like briefly to survey most relevant previous solutions of SFS problem. Since the problem has very long history, the full overview is out of the scope of this work, but nice and comprehensive SFS techniques analysis can be found in Kimmel et al. (1995) and Zhang et al. (1994). The SFS problem was first analyzed in computer vision by Horn in the early 1970s (Horn, 1975). His first solution involved characteristic strips expansion of the irradiance equation (1) and was not stable in practice due the noise sensitivity and error accumulation problems. During 1980–1990 in his work with Ikeuchi and Brooks another approach emerged (see Ikeuchi and Horn, 1981; Horn and Brooks, 1986). Using the calculus of variations, Horn and his coworkers reformulated the problem and solved it by minimizing an integral functional of the form

$$E = \iint_D E(p,q,z)dx\,dy = \iint_D (E_B + E_S + E_I)dx\,dy$$

where

$$E_B = (I(x,y) - R(p,q))^2$$
$$E_S = \lambda(p_x^2 + p_y^2 + q_x^2 + q_y^2)$$
$$E_I = \mu((z_x - p)^2 + (z_y - q)^2)$$

We say that E_S and E_I are respectively smoothness and integrability regularization terms and μ and λ are scalar multipliers. Obviously, E_S is small whenever the surface does not vary too fast. The minimization was done by iterative approximation of the associated system of three Euler equations where p, q, and z are treated as independent variables. At each step the solution must be feasible, i.e., $p = z_x$ and $q = z_y$, so the third, so-called integrability constraint, is imposed. The scalar parameters μ and λ are representing the weight of the penalty terms and they have to be carefully tuned. Too large λ, for example, might cause oversmoothing of the solution surface and important details would be lost. Researches have also tried to work with different penalty terms, like e.g.

$$E_S = (R_x - I_x)^2 + (R_y - I_y)^2$$

in the work of Zheng and Chellappa (1991) or totally different approaches like in Hsieh et al. (1994), Jones and Taylor (1994), Kimmel and Bruckstein (1995), Pong et al. (1989), Ron and Peleg (1990), Szelisky (1991), and Wei and Hirzinger (1997). However, Horn's variational method had been the "main stream" approach to the SFS problem until mid of 1990s.

B. Model-Based SFS. Another important idea was suggested by Atick et al. in 1996 (see Atick et al., 1996). In this work the authors dealt with the problem of recovering of the shape of human faces from their shaded images. To employ the a priori information about the class of objects studied, Atick gives up generality and solves the SFS problem for a particular class of surfaces. Given the exact mappings in a cylindrical coordinate system, he regarded each mapping as an independent realization of some stochastic process. Thus a "head-surface" can be represented as

$$r(\theta, l) = r_0(\theta, l) + \sum_{i=1}^{M} a_i u_i(\theta, l)$$

where $r_0(\theta, l)$ is a "mean-head,"

$$r_0(\theta, l) = N^{-1} \sum_{i=1}^{N} r_i(\theta, l), \quad \{r_i(\theta, l)\}_{i=1}^{N} (\text{database})$$

and $u_i(\theta, l)$ are the first M components of Karhunen–Loeve decomposition derived from the scanned surfaces, so-called "eigenheads." So, denoting

$$\vec{\alpha} = (\alpha_1, \alpha_2, \ldots, \alpha_M)$$

the SFS problem can be reformulated as

$$\min_{\vec{\alpha}} \sum_{l,\theta} (I(x(l, \theta), y(l, \theta)) - R(p, q; \vec{\alpha}))^2$$

which has been solved using gradient descent method. Here, the SFS problem becomes one of optimal parameter estimation. The most important advantages of this approach are the following:

1. Since $u_i(\theta, l)$ are smooth and feasible functions, the whole $r(\theta, l)$ is smooth and feasible, i.e., there is no need in E_S and E_I penalty terms, and hence, no need in tuning of λ and μ parameters.
2. While solving Euler–Lagrange system one needs some boundary conditions. If the choice of these conditions is "bad," the solution may wander far from the optimal one. Here the "mean-head" is a natural choice to start the minimization process.

Note that this model-based surface representation is low-dimensional, which makes the numerically complex SFS problem much more tractable.

IV. THE NEW METHOD

A. A Priori Knowledge. Unlike the work of Atick mentioned above, with surfaces of wafers we cannot obtain any kind of statistical information, except, perhaps, the CAD design itself which may be seen as a "mean-wafer," but no meaningful representation of a perturbation about this "mean-wafer" is available. Therefore, we concentrate our efforts to employ a priori knowledge about VLSI wafer geometry in an attempt to obtain low-dimensional and meaningful representations of wafer surfaces. Wafer surfaces could be characterized as a plane with a set of mutually disjoint wells and slots (see Fig. 1). These slots might be either straight or bent (so-called "L-shaped" slots). We assume that all the slots are parallel

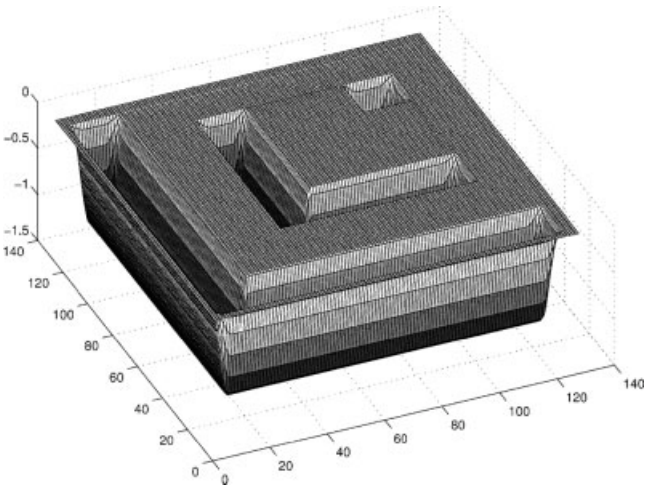


Figure 1. Surface model.

to the X and/or Y axes of the image plane. This assumption does not always hold in practice, but its violation is rather rare and hence disregarded in our model. The shape of wells may vary in their size. Wells of relatively large size look like a rectangle with somewhat rounded corners and edges that are normally parallel to the axes. However, smaller wells assumed to have rather rounded shape like ellipses or circles. One cannot assume that slopes of wells are symmetrical. Also no assumptions can be made about the small wells except for their rounded shape. In the next section we define mathematical objects that fit well the geometry described above.

B. The Mesa Functions. We start with the following formal construction of a so-called mesa (or table) function well known from the theory of distributions (Zemanian, 1965). Let us define

$$f(x; x_0, \varepsilon) = \begin{cases} c \exp\left(\dfrac{1}{1 - \left(\dfrac{x - x_0}{\varepsilon}\right)^2}\right), & |x - x_0| < \varepsilon \\ 0, & \text{elsewhere} \end{cases}$$

The constant c in this definition is such that

$$\int_{x_0 - \varepsilon}^{x_0 + \varepsilon} f(x; x_0, \varepsilon) dx = 1$$

Then we introduce an auxiliary function $\tilde{f}$ as following:

$$\tilde{f}(x; x_r, \varepsilon_r, x_l, \varepsilon_l) = f(x; x_l, \varepsilon_l) - f(x; x_r, \varepsilon_r)$$

and a mesa function is then defined as (see Fig. 2)

$$F(x; x_r, \varepsilon_r, x_l, \varepsilon_l) = \int_{-\infty}^{x} \tilde{f}(t; x_r, \varepsilon_r, x_l, \varepsilon_l) dt$$

The parameters of mesa function have simple geometrical interpretation:

1. $|x_r - x_l|$ is responsible for the size of the "pulse."
2. ε_r, ε_l control how fast the function climbs up and falls down.

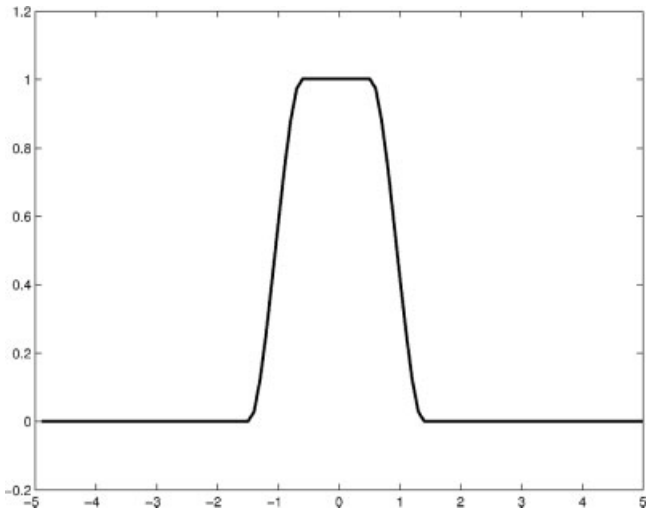


Figure 2. Mesa function.

Now, let us define 2D mesa function of height h to be h times the tensor product of two 1D mesa functions. For simplicity, we denote

$$\vec{\alpha}_x = (x_r, \varepsilon_{xr}, x_l, \varepsilon_{xl}); \quad \vec{\alpha}_y = (y_r, \varepsilon_{yr}, y_l, \varepsilon_{yl})$$

$$\vec{\alpha} = (x_r, \varepsilon_{xr}, x_l, \varepsilon_{xl}, y_r, \varepsilon_{yr}, y_l, \varepsilon_{yl}, h) \; = \; (\vec{\alpha}_x, \vec{\alpha}_y, h)$$

and then (see Fig. 3)

$$T(x, y; \vec{\alpha}) = h \cdot F(x; \vec{\alpha}_x) \cdot F(y; \vec{\alpha}_y)$$

C. Surface Representation. By means of 2D mesa functions we can build successful representations of surfaces of the type described in the previous section. Indeed, a straight slot feature can be represented as a negative mesa function with

$$|x_r - x_l| \ll |y_r - y_l|$$

or

$$|x_r - x_l| \gg |y_r - y_l|$$

To describe a well, we shall let $|x_r - x_l| \cong |y_r - y_l|$ and

$$\begin{cases} |(x_r - \varepsilon_{xr}) - (x_l + \varepsilon_{xl})| \ll \min(\varepsilon_{xr}, \varepsilon_{xl}) \\ |(y_r - \varepsilon_{yr}) - (y_l + \varepsilon_{yl})| \gg \min(\varepsilon_{yr}, \varepsilon_{yl}) \end{cases} \tag{2}$$

then the size of the "pulse" is negligible and the shape becomes rather elliptical. If condition (2) does not hold, we can characterize the shape as a "rectangle with rounded corners." An "L-shaped" feature can be constructed as difference between two 2D mesa functions when one of them is shifted from another. So we can write now a surface as a superposition of several 2D mesa functions, when each of them is fully defined by its vector of parameters.

$$z(x, y) = \sum_{i=1}^{N} T_i(x, y), \tag{3}$$

with

$$T_i(x, y) = T(x, y; \vec{\alpha}_i)$$

which means that in terms of this model, SFS problem becomes one of optimal parameter estimation. In the following chapters, we shall develop a technique to separate between the features and find N sets of parameters for each feature instead attempting to estimate them all together.

We would like to stress here two major difficulties that we have to cope with. A first and obvious one is that in formula (3) the number of terms in the sum N is, actually, unknown. Another weakness is the necessity for a good initial surface estimate to start the minimization process.

D. Strategy. In light of the problems presented above, let us briefly describe our main strategy. It is common practice to take into account prior knowledge about image structure (like edge-geometry or statistics of the noise) to simplify the very complex tasks of image processing and analysis. In our case, the surface model described above can be used to derive a set of simple properties of the resulting images, that are almost independently from the image formation rule. One can readily observe that the mesa function model induces a certain geometrical structure of the image. If the surface studied is a superposition of several mesa functions, then its image will have the following properties. First, all edges will be straight and parallel to the axes. Second, the edges are expected to be organized in well-defined constellations. For example, in case of an image of a large well, its edges will form a rectangular shape (see Fig. 4). Therefore, by decomposing the image in the Haar wavelets basis, and picking the largest amplitude coefficients by means of simple thresholding, we shall compact and effective representation of image singularities. Moreover, thanks to the similarity property of wavelet multilevel representation, large Haar coefficients will carry the geometrical information of image singularities at each level of the decomposition. Since image singularities form well-organized clusters, so do the large amplitude wavelet coefficients. An algorithm to discover these clusters can be developed. Then we shall exploit wavelets representation properties to localize every surface feature from its coefficients cluster and roughly calculate its

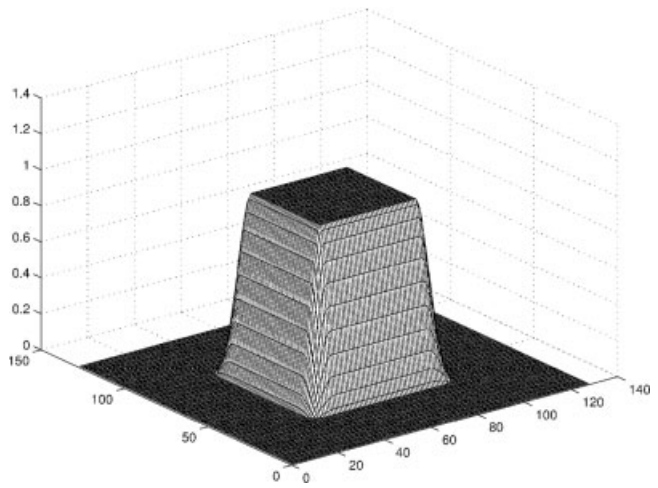


Figure 3. 2D Mesa function.

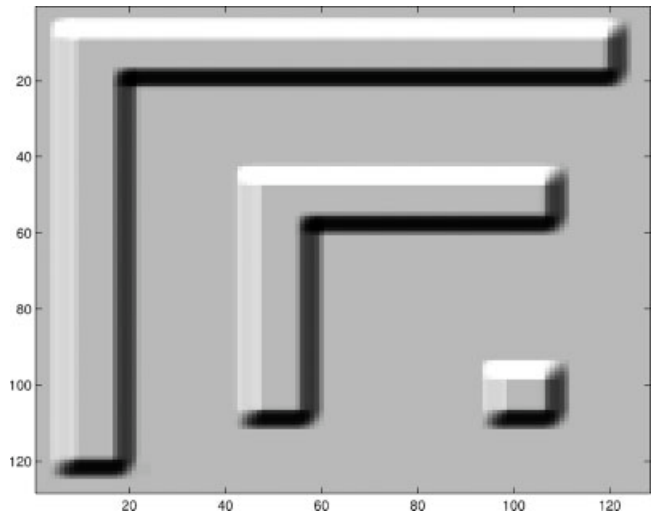


Figure 4. Lambertian image of the surface shown in Figure 1.

parameter vector. This vector will subsequently serve as the starting point for the least square error (LSE) optimization process and iteratively refined toward the final result. Two points are worth mentioning. First, applying this strategy, we will separate surface features and treat them independently, which solves the problem of unknown number of terms in model (3). Second, the initial parameter estimates are geometrically meaningful, which significantly reduces our chances to fall into some local minima during the LSE procedure. Let us elaborate on this strategy in the following sections.

E. Haar Wavelets. The theory of wavelets has had a major impact on the field of signal and image processing and computer vision. Wavelets and their possible applications have been extensively studied by many researches during the last decade (Mallat, 1999). Presenting the essentials of wavelet analysis is out of the scope of this paper; however, the definition of Haar Wavelet Transform, along with justification of its choice, has to be provided. In this section we follow the formalism introduced by Mallat (1999). Let us denote 1D Haar scaling and mother-wavelet functions respectively by $\Phi(x)$ and $\Psi(x)$, where $\Phi(x)$ is an indicator function of [0,1] and $\Psi(x) = \Phi(2x) - \Phi(2x-1)$. In 2D case, if we denote the three Haar mother-wavelets as

$$\Psi^1(x,y) = \Phi(x)\Psi(y),$$
$$\Psi^2(x,y) = \Psi(x)\Phi(y),$$
$$\Psi^3(x,y) = \Psi(x)\Psi(y)$$

If we define the dilations and translations of 2D Haar mother-wavelets as

$$\Psi^1_{j,n_1,n_2} = 2^{-j}\Psi^1(2^{-j}x - n_1, 2^{-j}y - n_2),$$
$$\Psi^2_{j.n_1,n_2} = 2^{-j}\Psi^2(2^{-j}x - n_1, 2^{-j}y - n_2),$$
$$\Psi^3_{j.n_1,n_2} = 2^{-j}\Psi^3(2^{-j}x - n_1, 2^{-j}y - n_2)$$

and denote $n = (n_1, n_2)$ then one can prove that the system of functions $(\Psi^1_{j,n}, \Psi^2_{j,n}, \Psi^3_{j,n})_{(j,n)\in Z^3}$ constitutes an orthonormal basis of $L^2(R^2)$, see Mallat (1999). The expansion of $f(x,y) \in L^2(R^2)$,

$$f(x,y) = \sum_{k=1}^{3}\sum_{j=-\infty}^{\infty}\sum_{n=-\infty}^{\infty}\langle f, \Psi^k_{j,n}\rangle\Psi^k_{j,n}$$

has an interpretation in terms of image details aggregation at all resolution levels that range from 0 to $+\infty$. This plays a crucial role in our application. The inner products $\langle f, \Psi^1_{j,n}\rangle$ represent details in the horizontal direction, $\langle f, \Psi^2_{j,n}\rangle$ give the details in vertical direction and $\langle f, \Psi^3_{j,n}\rangle$ are the details in both directions (corners). Since we never relied on the orthonormality of the basis, other types of hierarchical subband decompositions can be effectively used. It is the suitable geometrical properties of Haar functions that make them appealing to our model. This is well illustrated by the decomposition exhibited in Figure 5.

F. Feature Tracking Graph. Our goal in this section is to define a structure, which will enable us to take advantage of the geometrical information carried by the large amplitude coefficients of the wavelet decomposition. We recall that at each resolution scale s three binary images H_s, V_s, and D_s, are produced. The images are built from the large amplitude coefficients corresponding to horizontal, vertical, and diagonal directions. For any binary image I, we denote $N(I)$ to be a number of connected components of I. Given the connected components of all three images

$$CC(H_s) = \{CC_j(H_s)\}_{j=1}^{N(H_s)}$$

$$CC(V_s) = \{CC_j(V_s)\}_{j=1}^{N(V_s)}$$

$$CC(D_s) = \{CC_j(D_s)\}_{j=1}^{N(D_s)}$$

let us define a feature tracking graph (FTG) G_s using the following definitions:

1. The set of vertices of G_s is equal to the set $CC(D_s)$.
2. The set of the edges of G_s will be divided to the two disjoint subsets named H-type and V-type edges as following:
   The unordered pair (v_p, v_q) of vertices is said to be H-type edge iff

$$\exists k : CC_p(D_s) \cap CC_k(H_s) \neq \{\phi\}$$

   and

$$CC_q(D_s) \cap CC_k(H_s) \neq \{\phi\}$$

   and it is said to be V-type edge iff

$$\exists k : CC_p(D_s) \cap CC_k(V_s) \neq \{\phi\}$$

   and

$$CC_q(D_s) \cap CC_k(V_s) \neq \{\phi\}$$

Examples of binarization and corresponding FTG are presented on Figures 6 and 7.

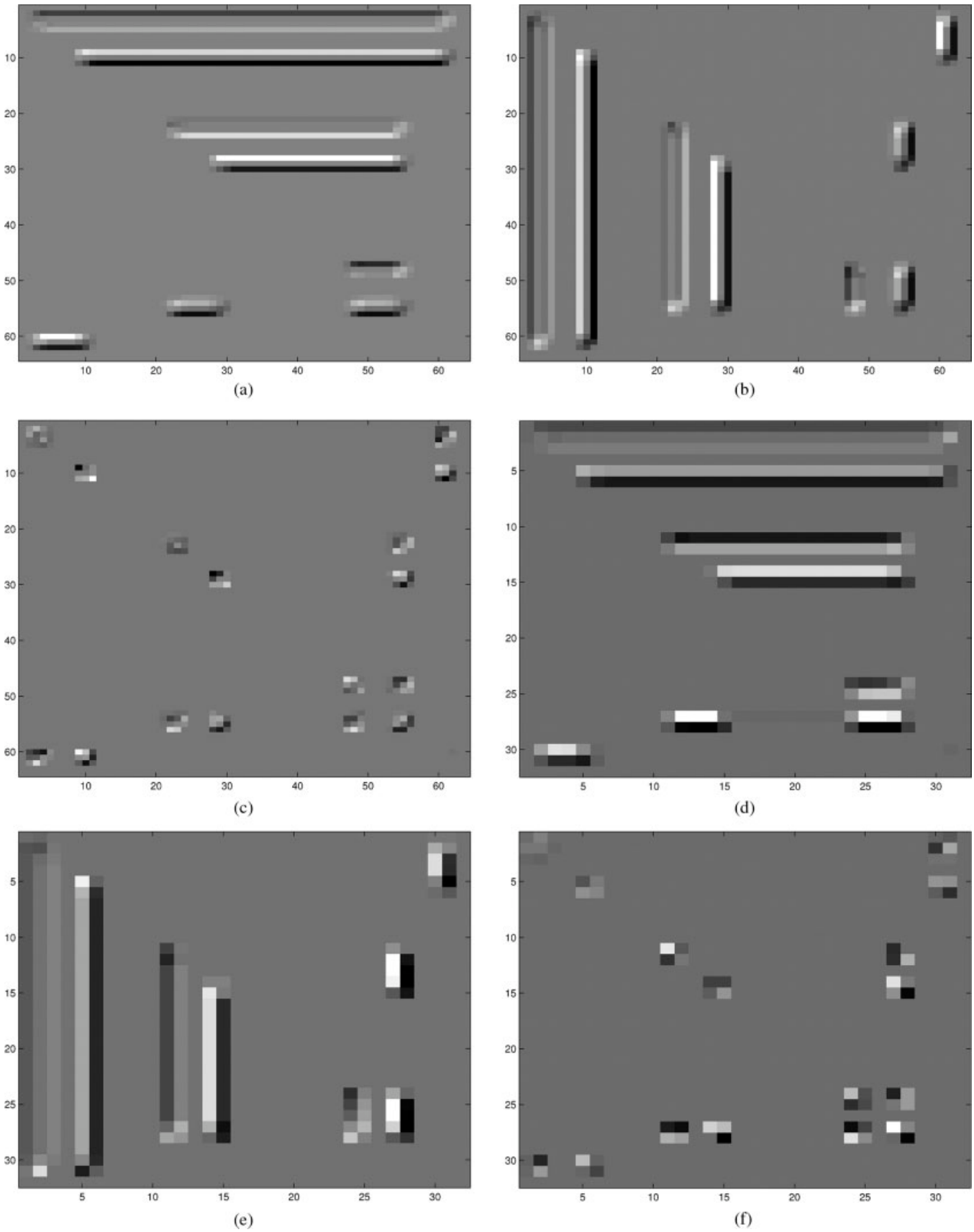


Figure 5. Haar wavelet decomposition of the image shown in Figure 4. White and Black dots respectively represent large positive and negative coefficients. Gray areas represent nearly zero coefficients.

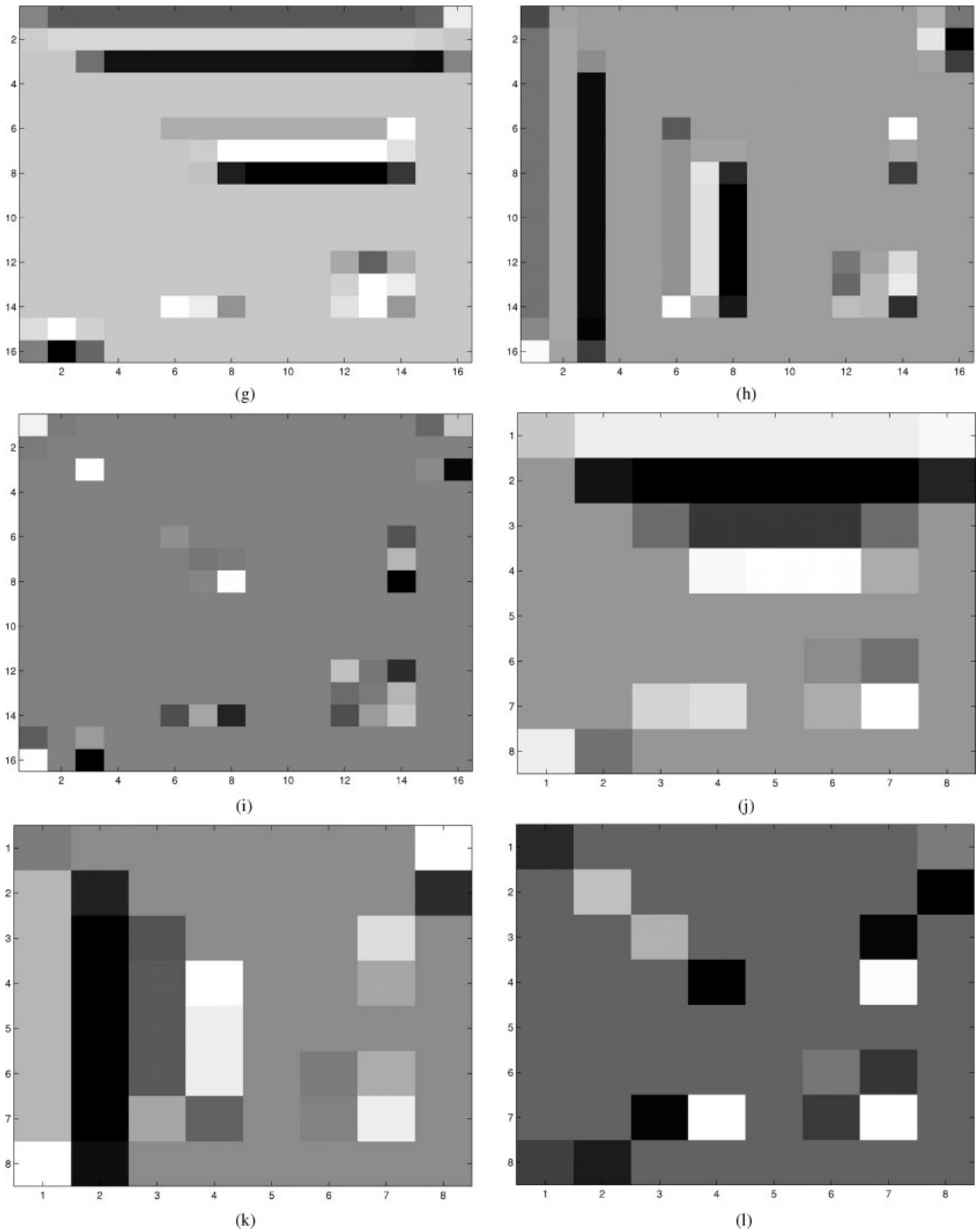
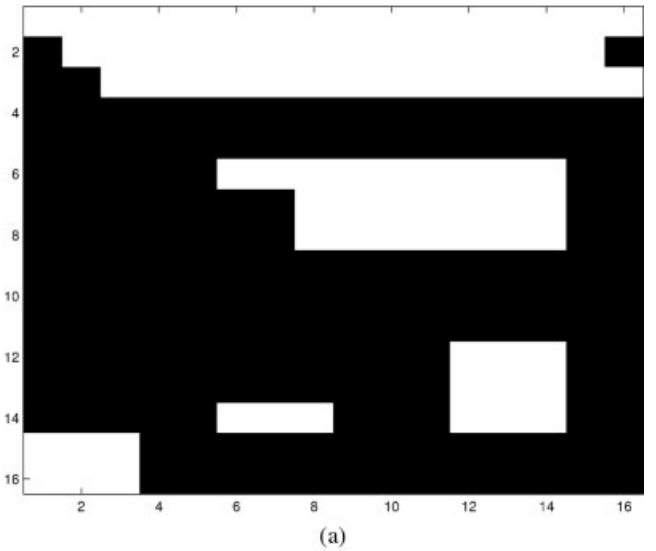
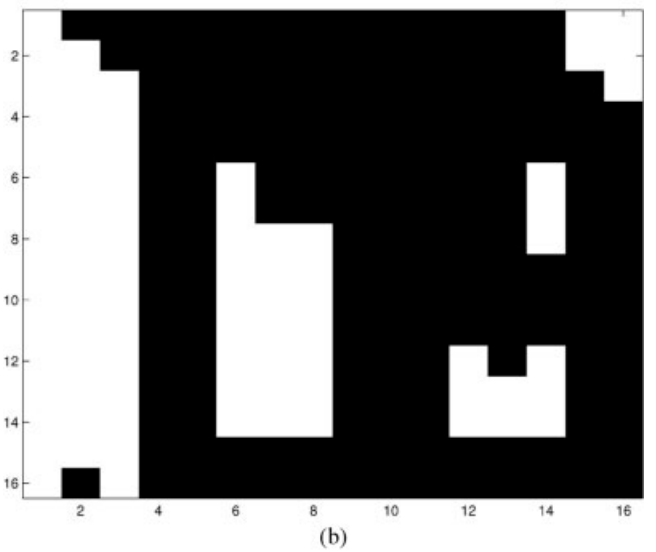


Figure 5. (*Continued*)





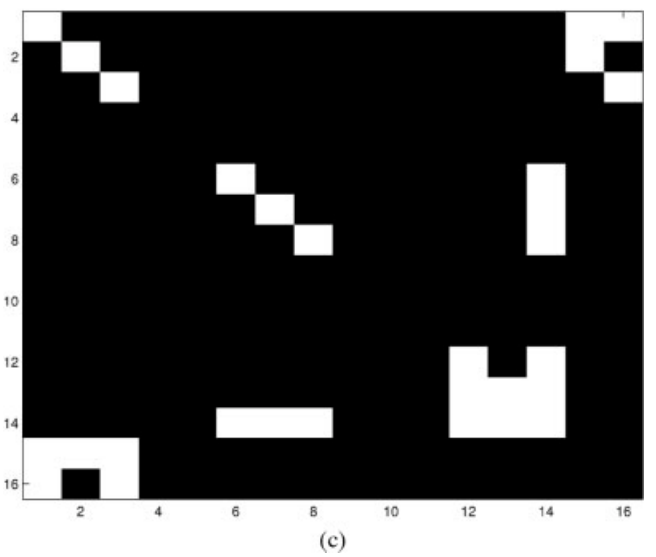


Figure 6. Binarization of the H, V, and D components of the Haar decomposition presented on Figures 5g, 5h, and 5j.

An unconnected FTG G_s can be decomposed to the set of N connected subgraphs

$$G_s = \{CSG_i(G_s)\}_{i=1}^{N}$$

by means of any standard method. The set of all pixels that constitute the connected subgraph CSG_i (G_s) will be denoted by P_i (G_s). We mentioned in Section IV.D that according to the model, if surface singularities are expected to be organized into certain geometrical structures, so are the constellations of the largest wavelet decomposition coefficients. Decomposition of FTG in connected subgraphs naturally reflects there geometrical structures at each level of resolution. For each CSG_i (G_s), using some properties of wavelet transform, it is possible to locate the spatial position of the surface singularities that originated the coefficients that CSG_i (G_s) is built from.

G. Spatial Orientation Trees, Regions of Interest (ROIs), and Initial Estimates. To trace the location of surface features, we need a notion of the spatial orientation trees (SOTs), borrowed from the works of Said and Pearlman (1996) and Shapiro (1993). In the works mentioned, SOTs have been used as a tool to optimize the transformation coefficients coding. Here we slightly modify their construction in order to adapt them to our goals. The Haar wavelet decomposition can be viewed as a hierarchical pyramid constructed with recursive four-subband splitting. On this pyramid, we define a quad-tree structure, which naturally reflects the spatial relationship between the coefficients on different scale levels in the following way. For each spatial orientation $k = 1, 2, 3$, we create quad-trees by relating recursively each coefficient at the scale s and position (p,q), say w_s^k (p,q), to its four children at the next, finer, scale $s + 1$:

$$w_{s+1}^k(2p, 2q), \quad w_{s+1}^k(2p+1, 2q),$$
$$w_{s+1}^k(2p, 2q+1), \quad w_{s+1}^k(2p+1, 2q+1)$$

Usually, the branching rule of the coefficients at the most coarse scale, say $s = 0$, is different. Each coefficient w_0^0 (p,q) is associated

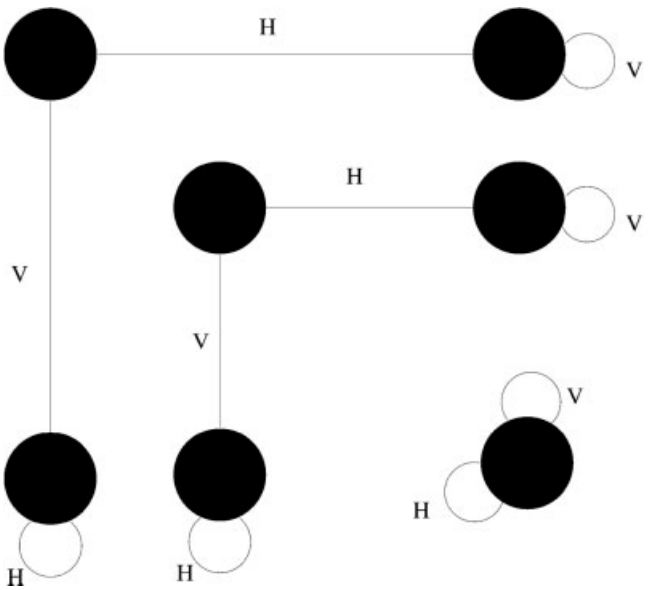


Figure 7. FTG built from the images in Figure 6.

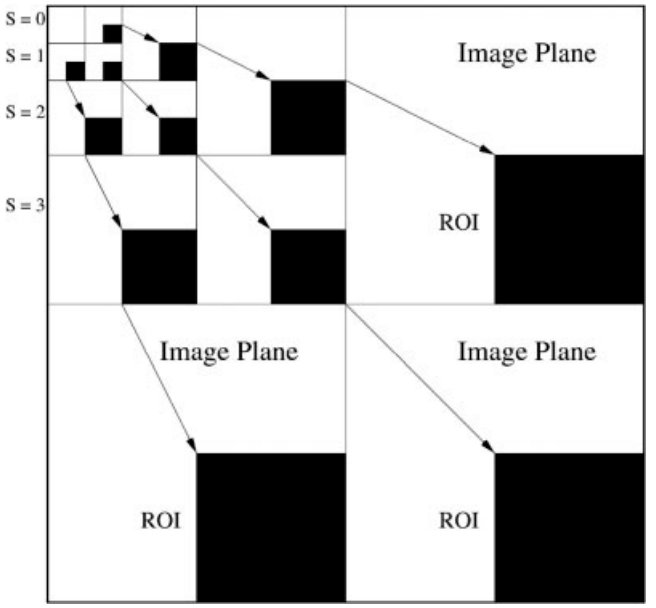


Figure 8. SOT construction and ROI.

with the three wavelet coefficients at the same scale and location: $w_0^1(p,q)$, $w_0^2(p,q)$, $w_0^3(p,q)$, and considered to be the roots of the trees. Unlike the classical SOT, we do not consider the coefficients $w_0^0(p,q)$ and build the trees from the coefficients of the scale level $s = 1$. In addition, we artificially attach the image planes to each spatial orientation. The reason behind it is that we are going to use not the values of the pixels, but only their absolute coordinates. The "pseudocoefficients" at this artificial level on the image plane will be considered as the leaves of the trees. The construction of SOT is illustrated in Figure 8. Note that if at some scale level s and position (p,q) there is a wavelet coefficient $w_s^k(p,q)$ of high amplitude, then all image singularities that possibly contributed to the coefficient are included into the spatial area of the leaves of the SOT rooted at (s,p,q). Let us denote it as the region of influence of $w_s^k(s,p,q)$ by ROI (s,p,q) (see Fig. 8). Clearly, the wavelet coefficients $w_s^1(p,q)$, $w_s^2(p,q)$, and $w_s^3(p,q)$ have the same ROI. Now, the ROI of each

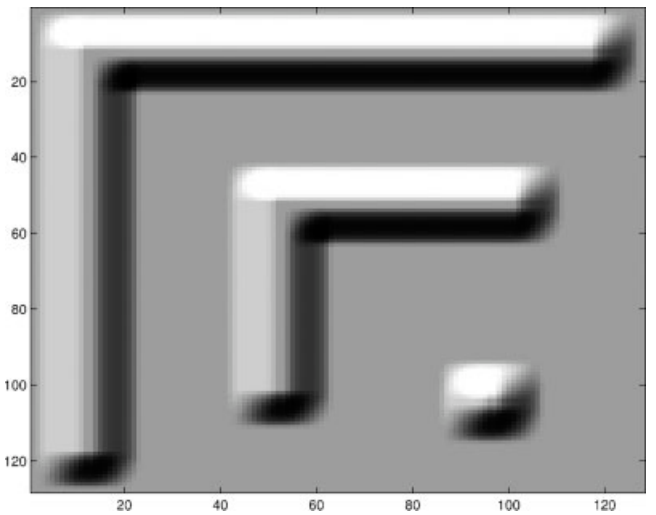


Figure 9. Initial guess image.

connected subgraph can be defined as

$$\mathrm{ROI}(CSG_i(G_s)) = \bigcup_{(p,q)\in P_i(G_S)} \mathrm{ROI}(s,p,q)$$

One can also see that by construction of the SOT, the ROIs of two different points (s,p_1,q_1) and (s,p_2,q_2) are disjoint.

$$\mathrm{ROI}(s,p_1,q_1)\bigcap \mathrm{ROI}(s,p_2,q_2) = \{\phi\}$$

and hence so are the ROIs of two different connected subgraphs:

$$\mathrm{ROI}(CSG_i(G_S))\bigcap \mathrm{ROI}(CSG_j(G_s)) = \{\phi\}, \quad i \neq j$$

Thus, the decomposition of the graph G_S to the mutually disjoint connected subgraphs enables us to define a set of mutually disjoint domains on the image plane where the potential surface features are located. Then, it is possible to calculate the set of the initial guesses for the feature location in order to initialize the optimization process.

We use a rather simple strategy to establish the initial estimates for the surface parameters. Given the sets of mutually disjoint connected subgraphs $\{CSG_i(G_s)\}_i$ and their ROIs $\{\mathrm{ROI}(CSG_i(G_s))\}_i$, we apply the following two-step heuristic decision:

Step 1: Decide on the "type" of the feature.

Step 2: Given the ROI, try to compute the parameters of a symmetrical feature of an appropriate type such that it would fully occupy the ROI (see an example in Fig. 9 and Table I).

The only parameter which does not play any role in this heuristic decision is the parameter "h" of the height of the mesa-function. We suggest to overcome this problem by choosing the initial value of the parameter according to the available data from the CAD of the wafer. By applying this scheme, two essential advantages can be achieved. First, since the ROIs are mutually disjoint so are the initial estimates and then the optimization processes will not interfere. This enables us to carry out the optimization computation in parallel, if appropriate hardware is available. Second, the optimization process per feature will run only on its own ROI and not on the whole image. This may provide considerable speed-up in such a heavy computational task as SFS.

H. The SFS Algorithm. We list below a summary of the steps of our new SFS algorithm.

Input

1. $I(x,y)$—the $M \times M$ gray level image of the studied wafer surface.
2. S—the coarsest scale level of the discrete wavelet transform.
3. LET—the local error tolerance value.

Table I. Initial guess vector.

x_r	x_l	ε_{xr}	ε_{xl}	y_r	y_l	ε_{yr}	ε_{yl}	h
9.10	−9.10	0.89	0.89	9.10	−9.10	0.89	0.89	−1.00
9.10	−7.14	0.89	0.89	9.10	−7.14	0.89	0.89	1.00
6.60	−2.69	0.89	0.89	6.60	−2.69	0.89	0.89	−1.00
6.60	0.89	0.89	0.89	6.60	−0.89	0.89	0.89	1.00
5.97	4.17	0.89	0.89	7.22	5.42	0.89	0.89	−1.00

Table II. Set of parameters of the test surface.

x_r	x_l	ε_{xr}	ε_{xl}	y_r	y_l	ε_{yr}	ε_{yl}	h
9.0	−9.00	0.5	0.5	9.0	−9.00	0.5	0.5	−1.00
9.0	−7.00	0.5	0.5	9.0	−7.00	0.5	0.5	1.00
7.0	−3.00	0.5	0.5	7.0	−3.00	0.5	0.5	−1.00
7.0	−1.00	0.5	0.5	7.0	−1.00	0.5	0.5	1.00
7.0	5.00	0.5	0.5	7.0	5.00	0.5	0.5	−1.00

Output

1. Number of features found
2. A vector of parameters of each feature
3. A vector of local errors induced by the calculated vector of the parameters

Step 1

1.1. Compute the Haar Wavelet Transform of the input image $I(x,y)$ up to the predefined scale level S.
1.2. Set the current scale level S

Step 2

2.1. Produce three binary images H, V, and D by picking the large amplitude coefficients of the corresponding parts of the pyramid at the current scale level s.
2.2. If the images H, B, and D are identically equal to zero—STOP.

Step 3

3.1. Compute the connected components of H, V, and D.
3.2. Build the graph G_s.
3.3. Compute the set of connected subgraphs $\{\mathrm{CSG}_i(G_S)\}_{i=1}^{N}$

Step 4: For each connected subgraph do

4.1. Decide on its type.
4.2. Calculate ROI ($\mathrm{CSG}_i(G_s)$).
4.3. Calculate the initial estimate vector $\vec{\alpha}_0$.
4.4. Launch a Levenberg Marquardt (LM) Optimization process on the ROI starting with the calculated initial estimate $\vec{\alpha}_0$, to determine

$$\min_{\vec{\alpha}} \sum_{(x,y)\in \mathrm{ROI}} (I(x,y) - R(x,y;\vec{\alpha}))^2$$

4.5. Save the resulting vector of parameters and the error introduced by it.

Table III. Set of reconstructed parameters.

x_r	x_l	ε_{xr}	ε_{xl}	y_r	y_l	ε_{yr}	ε_{yl}	h
9.0	−9.00	0.51	0.51	9.02	−9.02	0.51	0.51	−1.01
9.0	−7.00	0.51	0.51	9.02	−7.02	0.51	0.51	1.01
7.0	−3.00	0.50	0.50	6.99	−3.00	0.50	0.49	−1.00
7.0	−1.00	0.50	0.50	6.99	−1.00	0.50	0.50	1.00
7.0	5.00	0.50	0.50	6.99	5.00	0.50	0.50	−0.99

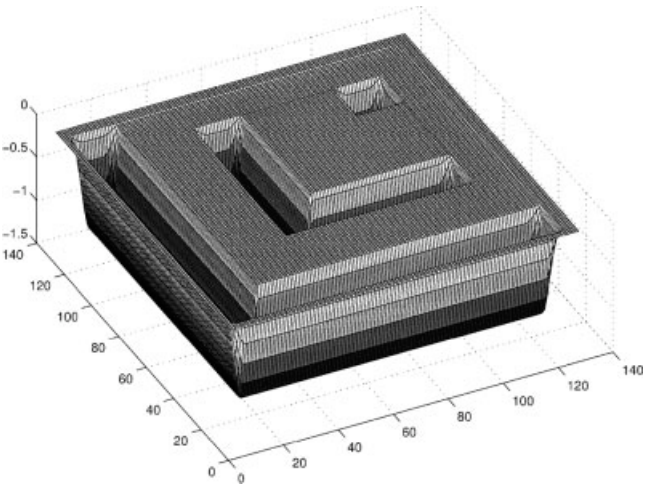


Figure 10. Surface reconstructed by our method.

4.6. If the error is less then the LET
4.6.1. For each pixel $(p,q) \in P_i(G_s)$ set all the wavelet coefficients in the SOT rooted at (s,p,q) to zero.
4.6.2. Output the resulting vector and the error.

Step 5
5.1. Set $s = s + 1$.
5.2. If $s > S$—STOP, otherwise GOTO Step 2.

It is worth mentioning that LM method was chosen here to address the ill-posedness of SFS problem and to insure stable convergence. The LM scheme is one of the most widely used nonlinear data fitting methods, and it is often described in the literature, see e.g., Press et al. (1988) and Gill et al. (1982). Since a regularization property is among its virtues, the method seems to perfectly fit our goal.

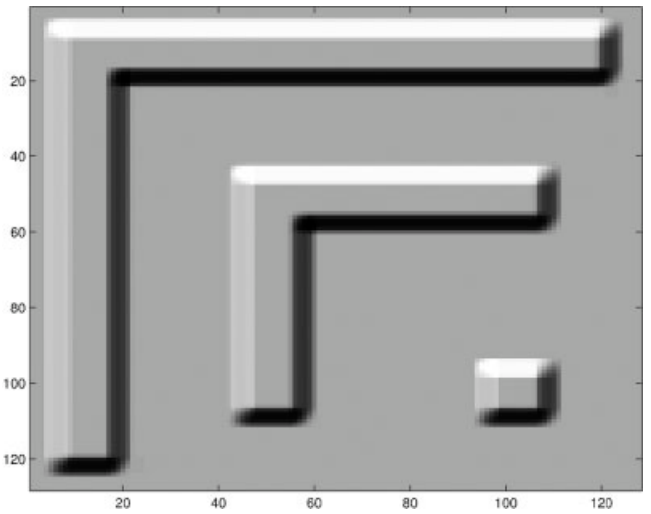


Figure 11. Image of the surface shown in Figure 10.

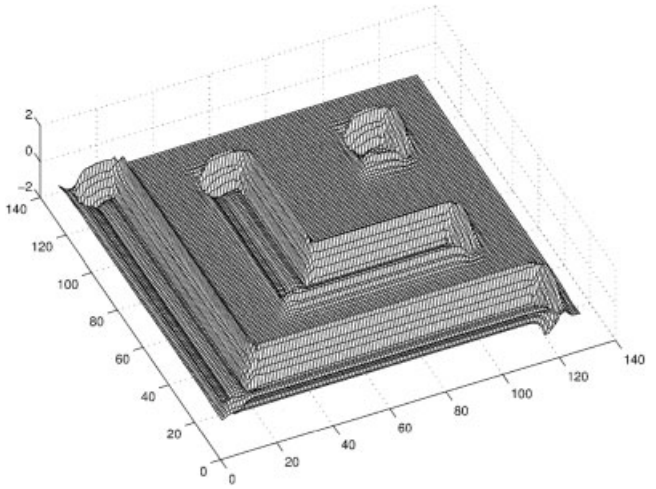


Figure 12. Surface reconstructed using Zheng and Chellappa method.

V. SUMMARY

A. Discussion. There are several important characteristics of the algorithm that we shall subsequently point out. First, a rather obvious observation is that the algorithm proceeds from the coarse to the fine scale levels. This does not mean that our algorithm "prefers" smoother solutions, like it was in the classical works of Ikeuchi and Horn (1981) and Zheng and Chellappa (1991), because we have not introduced any smoothness departure penalty term. However, the "coarse-to-fine" nature of the algorithm has some implications on its behavior. As we build the pyramid from the fine to the coarse scale levels, the "blobs" of high amplitude coefficients tend to mix together, meaning that the geometrical information carried by the "blobs" (the potential connected components and the elements of the FTG) is much more precise on the fine levels of the pyramid than on the coarse ones. Actually, if the image is absolutely free of noise, it is sufficient to choose $S = 1$ to recover the surface. However, in the presence of additive white noise, no comprehensible information will be contained at the finest level of the pyramid. Another conclusion is that by processing the pyramid from a coarse to a fine scale the algorithm will first recover the large and well-isolated surface features because the geometrical information about them is likely to "survive" at these levels. The dense constellations of smaller features can be recovered only from fine levels because their geometrical information will be lost at coarse levels of the pyramid. Note that if some "bold" feature has been recovered from some coarse level s of the pyramid and some connected subgraph CSG_i (G_s), then there is no point in taking into account the wavelet coefficient from finer levels belonging to the SOTs rooted at $(p,q) \in P_i$ (G_s). That is why the auxiliary Step 4.6.1. has been introduced.

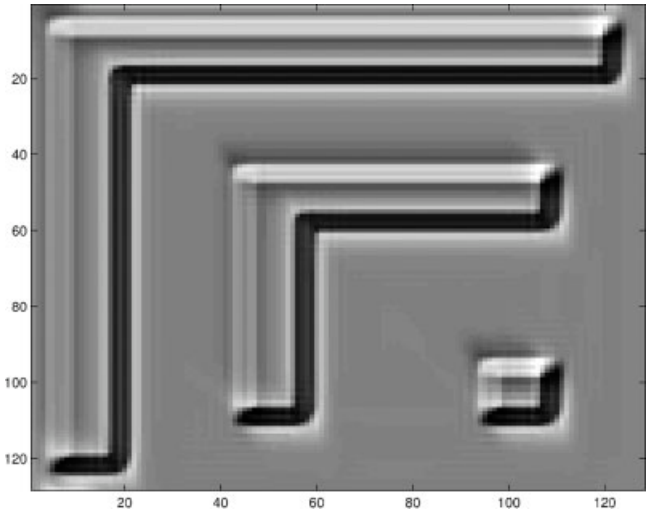


Figure 13. Image of the surface reconstructed using Zheng and Chellappa method.

B. Experiments and Results. To demonstrate the efficiency and accuracy of the proposed algorithm several synthetic images, both with and without additive Gaussian noise, have been selected for testing and comparisons. The algorithm has been applied under both Lambertian and SEM image formation models. The SFS algorithm proposed by Zheng and Chellappa (1991) has been implemented to provide data for the performance comparisons. One tested surface is shown in the Figure 1. It was constructed with the set of parameters shown in Table II. Its shaded Lambertian image is presented in Figure 4. Table III shows the resulting vectors computed by our algorithm, followed by the plot of reconstructed surface and its Lambertian image in Figures 10 and 11. Figures 12 and 13 show the surface and its Lambertian image reconstructed by the method of Zheng and Chellappa.

C. Conclusions. We have presented a new approach to the SFS problem. We employed a priori knowledge to develop a simple geometrical model for the class of possible surfaces and the SFS problem was reformulated in terms of a nonlinear data fitting problem. Wavelet decomposition has been used to determine an "intelligent initial guess" to start the minimization procedure with. The standard LM minimization algorithm, known to be stable and having a regularization property, was then adopted in order to produce an optimal solution.

ACKNOWLEDGMENTS

This research was partly supported by a grant from the Applied Materials Research Fund.

REFERENCES

J.J. Atick, P.A. Griffin, and A.N. Redlich, Statistical approach to shape from shading, reconstruction of three-dimensional face surface from single two-dimensional images, Neural Comput 8 (1996), 1321–1340.

P. Gill, W. Murray, and M. Wright, Practical optimization, Academic Press, London, 1982.

B.K. Horn, "Obtaining shape from shading information," In P.H. Winston (Editor), Psychology of computer vision, McGraw-Hill, New York, 1975, pp. 115–155.

B. Horn and M. Brooks, The variational approach to shape from shading, Comput Vis Graph Image Process 33 (1986), 174–208.

J. Hsieh, H. Liao, M. Ko, and K. Fan, "Wavelet-based shape from shading," International Conference on Image Process, Austin, Texas, 13–16 November, 1994, pp. 125–129.

K. Ikeuchi and B. Horn, Numerical shape from shading and occluding boundaries, Artif Intell 17 (1981), 141–184.

A. Jones and G. Taylor, Robust shape from shading, Image Vis Comput 12 (1994), 411–421.

R. Kimmel and A. M. Bruckstein, Tracking level sets by level sets: A method for solving the shape from shading problem, Comput Vis Image Understand 62 (1995), 47–58.

R. Kimmel, K. Siddiqi, B. Kimia, and A. M. Bruckstein, Shape from shading: Level set propagatioin and viscosity solutions, Int J Comput Vis 16 (1995), 107–133.

S. Mallat, A wavelet tour of signal processing, 2nd Ed., Academic Press, London, 1999.

D. Marr, Vision, W.H. Freeman, New York, 1982.

T. Pong, R. Haralick, and L. Shapiro, Shape from shading using the facet model, Pattern Recognit 22 (1989), 683–695.

W. Press, S. Teukolsky, W. Vefferling, and B. Flannery, Numerical recipes in C: The art of scientific computing, 2nd Ed., Cambridge University Press, Cambridge, 1988.

L. Reimer, Image formation in low-voltage scanning electron microscopy, SPIE Press, Bellingham, Washington, USA, 1993.

G. Ron and S. Peleg, Non-linear multiresolution. A shape from shading example, IEEE Trans Pattern Anal Mach Intell 12 (1990), 1206–1210.

A. Said and A. Pearlman, A new, fast and efficient image codec based on set partitioning in hierarchical trees, IEEE Trans Circuits Syst Video Technol 6 (1996), 243–250.

J. Shapiro, Embedding image coding using zerotrees of wavelet coefficients, IEEE Trans Signal Process 41 (1993), 3445–3462.

R. Szelisky, Fast shape from shading, CVGIP: Image Understanding 53 (1991), 129–153.

G. Wei and G. Hirzinger, Parametric shape from shading by radial basis functions, IEEE Trans Pattern Anal Mach Intell 19 (1997), 353–365.

A. H. Zemanian, Distribution theory and transform analysis, McGraw-Hill, New York, 1965.

R. Zhang, P. Tsai, J. Cryer, and M. Shah, Analysis of shape from shading techniques. ICIP, 1994, pp. 377–384.

Q. Zheng and R. Chellappa, Estimation of illuminant direction, albedo and shape from shading, IEEE Trans Pattern Anal Mach Intell 13 (1991), 680–702.