Algorithms for the Computer Simulation of 2D Projections from Structures Determined by Dividing Surfaces

David M. Anderson
Department of Biomaterials
University of Buffalo
Buffalo, NY 14214

Jayesh Bellare
Department of Chemical Engineering
Indian Institute of Technology
Bombay 400 076 India

James T. Hoffman and David Hoffman
Department of Mathematics
University of Massachusetts
Amherst, MA 01003

Janelle Gunther and Edwin L. Thomas
Department of Materials Science and Engineering
MIT
Cambridge, MA 02139

Abstract

Transmission electron microscopy (TEM) data come in the form of two-dimensional projections of specimens with finite width. For visual interpretation of such data it is valuable, and for quantitative comparison with models essential, to have available a method for generating simulated projections from model structures. We present three computer algorithms that simulate TEM micrographs, from structures determined by contrast between two materials separated by complicated dividing surfaces. Three very general forms used to represent dividing surfaces are treated: traditional coordinate systems, finite element representations, and tesselations adapted to complex line integration schemes. Because surface tension drives dividing surfaces to minimize their area, particular attention is paid to minimal surfaces, surfaces of nonzero constant mean curvature, and to parallel surfaces that may form in specimens constrained to small dimensions. Besides the model structure itself, the other inputs are the orientation of the model structure, the location of (parallel) upper and lower truncating planes representing the finite specimen thickness or the form of the bounding surface for microdroplet samples, and the direction and magnitude of a linear deformation representing distortion due to the microtoming process. The values of these registration parameters providing the optimal fit with digitized TEM data are found by a relaxation method. Remarkable matches are obtained between micrographs of block copolymer morphologies and model structures determined by surfaces of constant mean curvature.

1 Introduction

Most of the objects we encounter in daily life are opaque. We see their outer surfaces only, while their interiors as well as objects occluded by them are invisible to us. If an analogy were drawn between our eyes and an electron microscope, our vision is more akin to SEM (scanning electron microscopy) than to TEM (transmission electron microscopy). It is only natural, then, that we have difficulty deciphering TEM micrographs, which are two-dimensional projections through (often complex) three-dimensional objections that are translucent to the electron beam.

To aid in the interpretation of TEM data, computer algorithms have been developed, and these fall into two general classes. In the first, tilting capabilities of TEM microscopes are used to obtain projections of the three-dimensional object from many viewing directions. Since the Fourier transform of each projection corresponds to a central section of the three-dimensional Fourier transform of the object, direct Fourier methods can be used to reconstruct the object [1]. Methods for the restoration of the reconstructed objects attempting to account for incomplete information, in particular the missing cone in reciprocal space due to tilting limitations, include the Gerchberg-Papoulis method [2]; [3], and the more recent method known as the method of projections on to convex sets [4].

In the second class of algorithms, competing model structures are evaluated by computing projections and comparing the simulations with actual TEM data. The MULTISLICE program, developed by the Arizona group [5], is an example of such a program. MULTISLICE computes projections of crystals for high-resolution microscopy simulation, accounting for microscope optical parameters. MULTISLICE is intended for atomic-level simulations, so that objects are defined in terms of atoms.

This atomic description is not sufficient, however, in supramolecular structures. In particular, interest in our group often focuses on microphase segregated structures with characteristic length scales on the order of 10-100nm. In one case of a star diblock copolymer with a complicated microdomain morphology [6], a crude model was concocted using MULTISLICE, by placing spheres in strings to simulate tubular channels. The authors are not aware of any TEM projection simulation program designed for such microstructures.

In this paper we present several algorithms specifically designed to handle complicated microstructures, and demonstrate in addition a method for finding the optimal match to a TEM micrograph using a given three-dimensional model. Rather than imposing a shape, such as spheres, to the building blocks of the model structures, we allow a much more general form for the models: dividing surfaces are specified, which separate regions of different atomic potential. A specimen can be discretized into volume elements (voxels) which are filled with local atomic potentials as determined by the locations of the dividing surfaces. Alternatively, a specimen can be represented by discretizing its dividing surfaces into surface elements. The representation of structures with dividing surfaces is efficient—to obtain a resolution of δx, the data storage necessary goes as $1/(\delta x)^2$, rather than $1/(\delta x)^3$, as for voxel densities. Because interfacial tensions tend to minimize the areas of dividing surfaces, we place particular emphasis here on surfaces of constant mean curvature and minimal surfaces, which minimize area under volume fraction constraints and fixed boundary conditions, respectively.

In Section 2, we describe the projection algorithms. Section 2.1 describes an elementary scheme wherein the specimen is divided into voxels. Section 2.2 describes a relatively simple "brute force" scheme, in which a very large number of points are tested to see which material they lie in. The method is illustrated for a triply-periodic surface represented by finite elements, and for a doubly-periodic surface with a representation in Cartesian coordinates. In general the method will work with any surface division given in terms of a triangulation of the dividing surface, assumed to be periodic. Section 2.3 describes a more sophisticated scheme, which is an adaptation of a ray-tracing program. The method is adapted to surfaces that have been computed by a complex line integration known in the theory of minimal surfaces as the Weierstrass representation. In Section 3 a relaxation scheme is described which searches for the optimal fit of the model with digitized TEM data, by adjusting the parameters specifying the orientation, scaling, truncation and deformation of the model.

In Section 4 we present results of the algorithms and demonstrate some remarkable matches between TEM data and simulated projections.

2 Methods

In the methods to be described in this section, we assume that we are given a digitized electron micrograph, in the form of an $N_0 \times N_0$ array E of experimental pixel intensities. Typically this would be obtained by a microdensitometer scan of an electron micrograph, with one pixel corresponding to a square region of linear dimension on the order of 50 microns, and with N_0 typically 128 or 256. The user selects a region of interest D within E of size $N \times N$, with $N \leq N_0$. We wish to produce an $N \times N$ array T of theoretical pixel intensities by a computer simulation, based on a structural model determined by a dividing surface S. The accuracy of the fit between the experimental array E and the theoretical array T can be quantified by a root-mean-square norm over all pixels in the region of interest; an appropriate definition of such a norm is described in Section 3.

2.1 Voxel method

2.1.1 Defining the specimen

In this method, the specimen is divided into N^3 voxels, $V_{i,j,k}, 1 \leq i,j,k \leq N$, and each voxel is assigned the local atomic potential of the specimen there. The assignment may come naturally from a three dimensional finite element solution where there are no obvious dividing surfaces, or the assignment may be made by scanning through the voxels based on an appropriate representation of the surfaces found in the specimen. An example of the scanning technique is illustrated below for a particular case of concentric shells—a set of parallel surfaces that is particularly well suited for the voxel method. Next, N^2 projection-rays $R_{i,j}, 1 \leq i,j \leq N$ are sent through the voxels so that the theoretical pixel intensities in the image array T are given by $T_{i,j} = \Sigma_k V_{i,j,k}$, i.e., the sum of the voxel atomic potentials along the projection direction. If the microscope specimen has been sectioned to a thickness smaller than that represented by the voxels, two planes M_1 and M_2 are used to truncate the voxels selected for projection. If the microscope specimen is tilted with respect to the basis in which the original voxels are defined, then the voxels are assigned to a new basis, rotated appropriately by a set of three user-specified Euler angles, before projection.

2.1.2 A case involving Dupin Cyclides

In physical systems that form parallel surfaces, e.g., lamellar surfactant or block-copolymer phases, the dividing surfaces that form depend on the types of defects allowed. One family of parallel surfaces is that of Dupin cyclides, whose defects are confocal ellipse and hyperbola. Dupin cyclides can be represented as in [7]:

$$r = r_e + \frac{r_h - r_e}{r_h - r_e}(R - er_e \cdot r_h)$$

where

$$r_e(\theta) = [\cos\theta, (1-e^2)^{1/2}\sin\theta, 0], \quad 0 \leq \theta \leq 2\pi$$

and

$$r_h(\phi) = [e\cosh\phi, 0, (1-e^2)^{1/2}\sinh\phi], \quad -\inf \leq \phi \leq \inf\,,$$

are the elliptical and hyperbolic defects. As the three parameters θ, ϕ and R are varied, the position vector r spans the entire specimen. Surfaces of constant R are parallel to each other and at a distance δR from each other. Therefore, the atomic position is conveniently specified as $\rho(x, y, z) = \rho(r(\theta, \phi, R))$ in the most general case, but usually as $\rho(R(\omega))$ where ω is a coordinate along R, the distance along the parallel layers. This yields atomic potentials whose level surfaces are parallel to parallel Dupin cyclides.

2.2 Projection-ray subdivision method

2.2.1 Defining the projection rays

In this method, projection-rays $R_{i,j}$, $1 \leq i, j \leq N$ are sent through the model structure and subdivided into very small subintervals. The midpoint of each subinterval is checked to determine which side of the dividing surface S it is on. These rays can be thought of as simulating paths taken by electrons, and the projection-plane P into which the rays are sent represents the photographic film on which the electrons record the image. Naturally, the direction $\vec{A} = (a_1, a_2, a_3)$ of the rays $R_{i,j}$ is normal to P, and without loss of generality we can take the equation of P to be $\vec{A} \cdot \vec{X} = 0$, with $|A| = 1$. In crystallographic terms, the a_i are the Miller indices of the viewing (or projection) direction.

The finite thickness of the specimen is represented by defining two planes M_1 and M_2, referred to as the "microtome planes", which truncate the model structure. Because goniometer-equipped electron microscopes have the capacity to tilt the specimen, the unit normal $\vec{N} = (n_1, n_2, n_3)$ to these microtome planes (taken to be parallel here, although this can easily be generalized) can be different from (a_1, a_2, a_3). The distance $d = |m_1 - m_2|$ between these planes $M_1 \colon N \cdot \vec{X} = m_1$ and $M_2 \colon N \cdot \vec{X} = m_2$ should coincide with the known specimen thickness, which is set in the actual microtoming process.

The projection-ray $R_{i,j}$ is parametrized by the variable s:

$$R_{i,j}(s) = s\vec{A} + r^0_{i,j}, \quad r^0_{i,j} = (x^0_1, x^0_2, x^0_3) \in P\,.$$

The truncation of the model structure by M_1 and M_2 means that on $R_{i,j}$, we need only check those points (X, Y, Z) satisfying $m_1 < N \cdot R_{i,j}(s) < m_2$, or equivalently, $s_1 < s < s_2$ where:

$$s_k = \frac{m_k - N \cdot R_{i,j}(s)}{A \cdot N}, \quad k = 1, 2\,.$$

The points r are chosen so that a square grid is formed on the projection plane P with mesh size δ. This can be done by choosing

$$r_{i,j} = \delta\left((i + i_0)\,\vec{b}'_1 + (j + j_0)\,\vec{b}'_2\right)\,,$$

where b'_1 and b'_2 form an orthonormal basis of P. The constants i_0 and j_0 are adjustable parameters in the fit of E with T; incrementing i_0 by 1, for example, would shift the simulated image T one pixel to the left.

In order to fit a model to data it is also necessary to rotate the simulated image to obtain registry. A rotation through the angle ϕ is easily accomplished by defining a new set of basis vectors $\{b_1, b_2\}$:

$$[b_1, b_2] = \begin{bmatrix} \cos\phi, & \sin\phi \\ -\sin\phi, & \cos\phi \end{bmatrix} [b'_1, b'_2]$$

Then the final form of the expression for the $r_{i,j}$ becomes:

$$r_{i,j} = (i + i_0 - i_1)\, b_1 + (j + j_0 - j_1)\, b_2\,,$$

where i_1 and j_1 are the indices of the lower left pixel in the selected region of interest in D; that is, the region of interest has the pixel range

$$(i_1, i_1 + N - 1) \times (j_1, j_1 + N - 1)\,.$$

This simplifies the process of choosing the adjustable parameters i_0 and j_0, in cases where the user may wish to select various regions of interest in E. The point is that as i_1 and j_1 are changed in order to window a different portion of E, registry can be maintained without changing i_0 and j_0.

In order to maintain registry throughout changes in the size ($N \times N$) of the windowed region in E, we define the mesh size δ in terms of N_0 rather than N. We have found that the most intuitive way to quantify the scaling of the simulated image is in terms of the number n_e of unit cell lengths per edge, in the full $N_0 \times N_0$ image. We thus make the definition $\delta = n_e/N_0$. The quantity n_e is in expressed terms of the unit of length in the model structure, which in the first case treated in this section will be the lattice parameter of a cubic lattice.

The remaining registry parameters involve the deformation that is intended to represent deformation of the sample due to specimen elongation during the microtoming process. The perfect, undeformed model structure will be subjected to a dilation or compression along a single direction. In general it would take two parameters to describe this direction, but we assume here that this direction lies *in the microtome plane*; that is, there is no distortion normal to the microtome plane. This comes from physical considerations, namely that the distortion should be expected to be in the direction of the motion of the knife. Thus we will define a new coordinate system C with the (x_1, x_2)-plane parallel to the microtome planes, and the x_1-axis in the direction of the deformation. This deformation will be described by multiplying the x_1-coordinate by $1 + \alpha$, where α, which we will call the strength of the deformation, can be positive, corresponding to elongation, or negative, corresponding to compression.

We begin by defining basis vectors of a reference coordinate system C', by choosing an oriented basis of $N^{\perp}$, say V_1', V_2', and letting $V_3 = N$. The direction of the deformation is defined by the angle η, between this deformation and V_1'. The desired coordinate system C has basis vectors given by:

$$[V_1 \mid V_2] = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} [V_1' \mid V_2']$$

In this coordinate system, the deformation is expressed by the simple replacement $x \to (1 + \alpha)x$.

One can compute that, in the C': $\{V_1', V_2', V_3'\}$ coordinate system, this deformation can be represented by the matrix:

$$\begin{pmatrix} 1+\alpha\cos^2\eta & \alpha\sin\eta\cos\eta \\ \alpha\sin\eta\cos\eta & 1+\alpha\sin^2\eta \end{pmatrix}.$$

The coordinate system C' is independent of the angle η.

It remains to give the conversion from X coordinates to the coordinates X' in the system C'. The x_3 coordinate is determined by the condition that the vector $X - x_3'V_3$ be orthogonal to v_3', i.e., that $x_3'V_3$ be in the same $N = v_3'$ plane as X.

With these coordinate transformations, we can convert any point (X, Y, Z) in the *deformed* structure to an equivalent point (x, y, z) in the perfect, undeformed structure. The coordinates (X, Y, Z) are first converted to the C' coordinates (x_1', x_2', x_3') as just described. The inverse of the matrix M is then applied to the vector (x_1', x_2'), yielding (x_1^+, x_2^+). And finally, these C' coordinates (x^+, y^+, z') are then converted to the final coordinates (x_1, x_2, x_3) in the undeformed structure via

$$(x_1, x_2, x_3) = a_1'x_1^+ + a_2'x_2^+ + a_3'x_3'\,.$$

The point (x_1, x_2, x_3) is then tested to determine which side of the perfect, undeformed dividing surface it lies on. This test, for the cases of finite element representations and traditional coordinate system parametrizations, is illustrated in the next two subsections. This will provide all the machinery needed to send projection-rays into the (deformed) structure and test an arbitrary number of points on each ray to determine which material they lie in. We have found that good results are obtained by dividing the interval $[s_1, s_2]$ into 50 subintervals, and testing the midpoint of each corresponding subinterval in the projection ray.

We have taken the optical density at a given pixel (i, j) to be a linear function of the fraction Ψ of subintervals of $R_{i,j}$ lying on one side of the intermaterial dividing surface (of given electron scattering power). This is because response curves of photographic films show that the optical density varies linearly (at least over a suitable range) with exposure, and although the mass-contrast mechanism by which our electron micrographs are produced is, strictly speaking, described by Beer's Law, the thickness of our specimens is small enough that the linear approximation to this exponential decay is sufficiently good. The two constants in this linear function are determined by fixing the mean and variance of T equal to those of E.

2.2.2 A case involving a finite element representation

We now describe, for a dividing surface with a finite element representation, the test for determining which side of the dividing surface a given point is on, i.e., which material it lies in. It has been shown [8] that within the framework of the finite element representation, there exists a tremendous amount of freedom in "customizing" generalized coordinate curves to match boundary conditions and to handle complicated geometries. The specific example we focus on here should illustrate how one can take advantage of this freedom. In particular, with any traditional coordinate system (x_1, x_2, x_3), the computational domain Q for any nonparametric representation $x_3 = x_3(x_1, x_2)$ of the surface would have to be determined along with the surface, whereas in the finite element representation, Q is the unit square, the most convenient computational domain possible. The same procedure, with minor modifications, has also been used [9, 10], for other model structures determined by surfaces of constant-mean-curvature ("H-surfaces") that were computed by one of the authors [11], using a finite element method.

The example, in which reflectional, translational, and rotational symmetries are all used, is the so-called "double-diamond" structure. This structure has recently been discovered experimentally in block copolymers [6]. It consists of two infinite, intertwined but distinct labyrinthine networks having diamond-cubic symmetry, which are filled with one material, e.g., polystyrene, while the matrix between them is filled with another material, e.g., polyisoprene. This matrix is bisected by an embedded (non-self-intersecting), triply-periodic minimal surface that was discovered in the 19th century [12], known as the "Schwarz diamond minimal surface". (A minimal surface is by definition a surface with zero mean curvature at every point.) The dividing surface S between the two materials is approximated by a surface of constant, nonzero mean curvature [9]. S is composed of two disjoint copies of an H-surface with $F43m$ symmetry, each of which surrounds one of the labyrinthine diamond-networks; this H-surface is shown in Figure 1. The two copies are congruent via a rotational symmetry. Thus, the space group $Pn3m$ of the full structure is of higher symmetry, due to this additional rotational symmetry.

The fundamental patch S_0 of the H-surface has been computed [11] using the following parametrization. As the surface coordinates (u, v) range over the unit square, and w ranges over the unit interval, the point

$$r(u,v) = \frac{1}{2}\left(1 - u(u+v)w,\ 1 - u + (u-v)w,\ u - 1 + (2 - u - v)w\right),$$

ranges over the body B, which is a tetrahedron known as a "tetragonal disphenoid"

[13]. This tetrahedron fills space without overlap on application of the reflectional symmetries of the space group. The surface patch S_0 is represented in (u,v,w)-coordinates as $w^*(u,v)$, where $w^* = w^*(u,v)$ is the solution of the finite element problem. A very simple scheme for arriving at such a parametrization for a given space group has been given [11].

The (u,v) unit square was divided into a 20×20 mesh—400 elements, 441 nodes—and 441 values w^* computed. The rotation that produces the copy of S_0 is given by:

$$x \to 1-x, \quad y \to y, \quad z \to -z\,.$$

This second copy also lies within B and is orthogonal to its faces.

An arbitrary point (x,y,z) in space can be reduced to an equivalent point inside B, that is, satisfying the conditions $0 \leq z \leq y \leq x \leq 1$, $x+y \leq 1$. The first step is to use the greatest integer function to truncate the integer parts of x, y, and z. Then the minimum and maximum functions are used to permute the three coordinates so that $z \leq y \leq x$ (the planes $x=y$ and $y=z$ are symmetry planes). If the condition $x+y \leq 1$ is satisfied, we are done. Otherwise we reflect across the plane $x+y=1$, using

$$x \to 1-y, \quad y \to y, \quad z \to z\,,$$

and the coordinates are permuted again. This is continued until all the conditions are satisfied. Let the coordinates of the final point be (x_0, y_0, z_0).

This point is converted to (u,v,w)-coordinates, using

$$w = 1 - x_0 + z_0, \quad v = (1-x_0-y_0)/(1-x_0+z_0), \quad u = (x_0-y_0)/(x_0-z_0)\,.$$

This triplet (u,v,w) is then tested to determine if $w > w^*(u,v)$. The solution $w^*(u,v)$ is interpolated from the values of w^* at the nodal positions using the finite element basis functions, in the standard way. If indeed $w > w^*$, then the original point (x,y,z) lies in one of the diamond labyrinths (e.g. in polystyrene).

If $w < w^*$, then another test must be performed to determine if (x,y,z) lies in the other diamond labyrinth. This is straightforward, because the rotation that interchanges the two labyrinths is given by:

$$u \to v, \quad v \to u, \quad w \to 1-w\,.$$

These new (u,v,w)-coordinates are tested as above to determine if $w > w^*(u,v)$.

2.2.3 A case involving a standard coordinate representation

In this subsection we describe the test in one case of a dividing surface with a representation in a standard coordinate system, in fact a Monge representation $z = z(x, y)$. The dividing surface (shown in Figure 2) is a minimal surface known as "Scherk's First Surface", after H. F. Scherk who discovered it in 1835 [14, 15]. This doubly-periodic minimal (zero mean curvature) surface can be qualitatively described as a family of evenly spaced parallel halfplanes above the (x_1, x_2)-plane that are smoothly joined to another family of halfplanes below the (x_1, x_2) plane, the families meeting at an acute angle γ, $0 < \gamma \le \frac{\pi}{2}$. The surface that is most familiar is the one with $\gamma = \frac{\pi}{2}$. In general, a surface in this family is given in implicit form as

$$F(x_1, x_2, x_3) =$$

$$e^{x_3} \cos(x_1/\alpha + x_2/\beta) \cos(x_1/\alpha + x_2/\beta) = 0\,, \quad \alpha^2 + \beta^2 = 1\,,$$

where $\alpha = \arccos \gamma$. The surface consists of a collection of graphs over a checkerboard of regions in the (x_1, x_2)-plane with vertical line boundaries at the vertices of the checkerboard.

To test whether a given point (x_1, x_2, x_3) is on the "A" side or the "B" side, we may simply check the sign of F. In practice, we do this as follows. First, define new x and y coordinates x' and y' in a coordinate system that follows the two asymptotic directions:

$$\begin{aligned} x' &= x/\cos\gamma \quad - y/\sin\gamma\,, \\ y' &= x/\cos\gamma \quad + y/\sin\gamma\,. \end{aligned}$$

We determine that the point is on the "A" side if either:

a) $\cos x' > 0$ and $\cos y' < 0$; or

b) $\cos x' > 0$ and $\cos y' > 0$ and $e^z < \cos x'/\cos y'$; or

c) $\cos x' < 0$ and $\cos y' < 0$ and $e^z > \cos x'/\cos y'$.

Otherwise the point is on the "B" side.

2.3 Ray-tracing method

2.3.1 Overview

This method works as in the previous section by first intersecting the projection ray R_{ij} with the two truncating planes M_1 and M_2 to obtain points b_1 and b_2 defining a line segment B, corresponding to the path of R through the specimen. B is then subdivided into a series of adjoining line segments C_k by intersecting it with the dividing surface S. Each component of C is on one side or the other of the dividing surface. The lengths of the segments on each of the two sides are summed and these two sums are used to compute a ratio giving the pixel intensity of the ray.

2.3.2 Surface Representation

The dividing surface S is represented as a data structure consisting of a mesh of triangular facets, which can be generated by any of a number of programs, including Mesh, Solsurf, and VPL applications [16]. Typically a triangulation of a fundamental piece of the surface is generated by Mesh then built up into a larger piece by a VPL application. The use of a piecewise approximation of the surface means that this algorithm imposes no constraints on the analytic properties of the surface.

The intersections of the line segments B_{ij} with the surface are found by computing the intersections of the lines with individual triangular facets. Without using any method to constrain the search for such intersections, the number of potential intersections that would have to be considered to compute an image would be equal to the number of facets in the model times the number of pixels in the image. Thus the cost of computing detailed images of accurate models could become prohibitive.

2.3.3 Space Partitioning Tree

To address the problem, this algorithm employs a data structure to partition space. By selectively accessing facets through this data structure, the number of facets that need to be considered as intersection candidates with a given line segment is drastically reduced. The data structure consists of a sparse binary tree of nodes, where each node occupies a rectangular volume. Nodes are classified as terminal and nonterminal. Non-terminal nodes have two descendents each, whereas terminal nodes have no descendents. A non-terminal node's descendents are are formed by bisecting its volume by a plane perpendicular to the X, Y, or Z axis. Each terminal node contains a list of references (pointers) to each facet having any part within the node's volume.

The execution of the algorithm consists of two phases: the construction of the tree of nodes given a set of facets, and the tree-constrained search for intersections of line segments with facets.

2.3.4 Tree Construction

The tree is built incrementally: An initially empty tree expands adaptively as new facet references are inserted in terminal nodes of the tree. The algorithm iterates over the list of facets, traversing the tree for each facet to determine what terminal nodes must have references to it. The tree is initialized as a single empty terminal node (the root node), and is allowed to grow as facet references are added. The tree is traversed from the root node for each facet. Traversal is accomplished by a routine that is called recursively on a per-node basis. It is called once at the root node for each facet. It tests whether the volume occupied by a node overlaps any portion of the facet. A node that does not overlap the facet is not considered further, and any descendents are ignored. A non-terminal node that overlaps the facet causes the routine to be called for the node's two descendents with the same facet. A terminal node that overlaps the facet adds a reference to the facet to its facet list. Since a facet may pass through the volumes of several terminal nodes, it may have several references. (See Figure 3.) A parameter sets the maximum number of facet references a terminal node may contain, and when this number is exceeded the node automatically bifurcates into a non-terminal node with two terminal descendents. Then the facet list of the defunct terminal node is processed to obtain the facet lists of the new terminal nodes. In this way the tree grows adaptively, acquiring greater depth and detail only in regions where the facets are dense.

2.3.5 Intersection Search

The set of all the facets intersecting the line segment B is found by searching the tree for all terminal nodes through whose volumes the line segment the passes. As in the case of facet insertion, searching is accomplished by a routine called recursively on a per-node basis, starting at the root node. It tests whether the volume occupied by a node overlaps a portion of B. A node that does not overlap B is not considered further, and any descendents are ignored. A non-terminal node that overlaps B causes the routine to be called for the node's two descendents. (See Figure 4.) A terminal node that overlaps the B may contain references to facets which intersect B. Each facet on its facet list is compared to the segment, and facets which intersect B within the volume of the referencing node are recorded on the list g. The restriction that the intersection lie within the node's volume is necessary to prevent redundantly recording intersections of facets referenced by more than one node.

2.3.6 Efficiency

The algorithm above provides an efficient method accessing the facets of a three-dimensional object with specified locii. The advantages of this method are especially apparent for highly complex objects, consisting of tens of thousands of facets. The space efficiency arises from the fact that the tree grows deep and resolves finely only those regions occupied by facets, spanning empty regions with relatively few nodes. Increasing the fineness of the mesh of facets approximating a surface, for example, will cause a less than proportional increase in the size of the tree. The time efficiency arises from the fact that a search for the terminal nodes occupying a small locus rapidly eliminates from consideration large portions of the tree, its speed being primarily dependent on the depth of the tree.

2.3.7 Intersection Processing

The list of intersections g is used to subdivide line segment B and to construct the series of linesegments C_k. The data associated with an intersection consists of its 3-dimensional spatial coordinates and its parity. The latter is determined from the parity of the facet, given that the facets are all created with a consistent parity (ie: a facet viewed from one side of the surface will have clockwise-ordered vertices, whereas one viewed from the other side will have counter-clockwise-ordered vertices.) The intersections in g are parameterized according to their position along B and sorted. Given that the mesh of facets approximating the dividing surface is properly generated, the sorted intersections will alternate in pairity. These intersections are used to cut B into the series of adjoining linesegments C_k. (See Figure 5.) If there is at least one intersection, the modes of the two or more components of C can be determined from the parity of the intersection(s). The modes of the components of C will, or course, alternate. If there is no intersection, the following method is used to determine the mode of the single component of C. A ray Q is constructed having its origin at b_1 and random direction. Any intersections between it and S are computed and recorded, and all but the closest intersection are discarded. If an intersection is found, its parity is used to determine the mode of b_1 and hence of C. If none is found, the process is repeated until an intersection occurs.

3 Matching experimental data

A measure of the discrepancy between the images T and D can be defined by the root-mean-square (rms) difference in the pixel intensities, normalized by dividing by the rms difference expected for a random T with the same mean and variance as D. If the standard deviation of the pixel intensities in D is σ, then this expected rms difference is simply $\sqrt{2}\sigma$. We have found that a value of $\varepsilon \leq 0.3$ for this normalized error ε corresponds to an excellent fit.

Having available a well-defined measure of the fit of T to D, we are in a position to carry out a relaxation method in order to determine the optimal registration parameters. Because the number of registration parameters is 14, the search over the 14-dimensional parameter space can easily lead to metastable minima. Therefore we have used the method of simulated annealing to facilitate the finding of the absolute minimum, although with a finite running time there is still no guarantee that the absolute minimum will be found.

In the method of simulated annealing, small random perturbations are made in the registration parameters, and at each change the increment δ in the normalized error ε is evaluated. If $\delta < 0$, then the change is retained; if $\delta > 0$, then the change is accepted with a Boltzmann-like probability $e^{-\mu\delta}$. The constant μ plays the role of an inverse temperature. We have obtained good results by setting μ equal to a small value in the initial stages of the fitting procedure, so that approximately half of the changes are accepted, and then incrementing μ so that in the final "fine tuning" stages only $5 - 10\%$ o the changes are accepted. In the case discussed in the results, the increment in the registration parameters was randomly distributed between about -0.02 and 0.02, although certain of the parameters such as the parameter n_e were perturbed by larger amounts. Only one parameter was changed at each step, and the value of μ was on the order of $10^3 - 5 \times 10^3$. On the order of 10^4 steps were taken at each stage, after which the value of μ was increased, and three or four such stages produced very satisfactory results. The initial estimates for the registration parameters, to start the fitting procedure, were obtained by trial-and-error.

In many electron microscopes it is possible to both rotate the sample in the plane P normal to the electron beam, and tilt it about a fixed axis F in P. If we choose a coordinate system in which the x_3-axis is parallel to the direction of the beam and the x_1-axis is in the direction of F before rotation, the relation between (ρ, τ) and the Miller indices (h, k, ℓ) is simply that $(1\rho, \tau)$ are the polar coordinates of (h, k, ℓ);

$$\begin{aligned} h &= \sin(\tau)\sin(\rho) \\ k &= \sin(\tau)\cos(\rho) \\ \ell &= \cos(\tau)\,. \end{aligned}$$

This transformation can be used in simulating a tilt series, which can provide evidence for a structural determination. In the simulation of a tilt series, the tilt angle determined by the relaxation method would be compared with the known tilt angle in the experiment.

4 Results

Microdroplet specimens can be produced by forming an aerosol of a diblock copolymer and volatile solvent. If a diblock copolymer which normally exhibits flat, parallel, equidistant surfaces ("lamellae") in the bulk state is employed, the microdroplet structure consists of a set of concentric spherical shells, of alternating composition with the lower surface tension polymer component (polyisoprene) in the outermost shell (see Figure 6a). Figure 6b illustrates a simulated image from such a simple specimen made with the voxel method. The voxel method was also used to produce the simulated image from a more complex set of parallel surfaces: Dupin cyclides (Figure 7). The images clearly show that the voxel method, though numerically intensive, yields revealing, characteristic images for specimens whose internal atomic potentials vary in prescribed manner.

In Figure 8a we show TEM data that have been digitized using a microdensitometer displayed on a computer screen and photographed. The image is of a star diblock copolymer which has been determined to microphase-segregate into the double-diamond microstructure discussed above [6]. In Figure 8b, the data of Figure 8a have been restored using a newly-developed maximum entropy algorithm [9] which maximizes the entropy of the power spectrum of the image, and thus restores and symmetrizes the data in a manner free from user-bias. Figure 8c, then, shows the best-fit theoretical projection T from the double-diamond model defined by the dividing surface of constant mean curvature described above. The match is striking, and the registration between D and T is made clear by Figure 8d, which is the left-half of Figure 8c juxtaposed with the right-half of Figure 8b.

The relaxation method described above was used to determine the registration parameters in this remarkable fit, and the final value of ε was 0.295. It is interesting that the value of α was about 0.22, thus indicating that the deformation due to microtoming was approximately 22%. Because small-angle X-ray scattering (SAXS) data on the same block copolymer indicated ($Pn3m$) cubic symmetry, this deformation must have occurred during the preparation of the sample for the electron microscope.

It should be mentioned that we have simulated projections of this model structure from many other viewing directions, and these have compared well experimental TEM data. In particular, we have obtained qualitative matches of the (100), (110), and (211) projections. However, only in the case of the (111) projection just shown have we used the relaxation method to obtain a quantitative fit with good registry. We have found this (111) projection to be the best signature of the double-diamond structure.

In Figure 9a we show a projection of the structure described in Section 2.2.3 based on the Scherk minimal surface with $\gamma = \frac{\pi}{2}$. The dividing surface in this case separates the $0s0_4$ stained electron dense regions of polyisoprene from the less dense polystyrene regions. The view is in the direction of the normal to one of the sets of asymptotical planes, so that on the upper part of the projection we see solid grey, corresponding to rays which pass alternately through the black and white domains. In the lower part we see alternating layers of black and white. Of key interest here is the "archway" appearance of the periodic saddle surface transition region which corresponds well with the experimental TEM data shown in Figure 9b.

Finally, Figure 10 is an approximation of the gyroid surface (space group Ia$\bar{3}$d) using the truncated Fourier series representation of a zero potential surface. Von Schnering and Nesper [17] have demonstrated the close relationship between certain periodic equipotential surfaces and certain triply periodic minimal surfaces. Barnes et al. [18] have recently shown that the equipotential surface is not a minimal surface nor is the truncated Fourier series representation of the equipotential surface. However, the simple trigonometric formulas provide a remarkable resemblance to certain minimal surfaces and moreover, provide a very efficient analytical means for surface visualization and subsequent projection simulation. The (100) and (111) projections of the approximate gyroid surface shown in Figures 10b and 10c were produced using ProDrop, a software package designed by Janelle Gunther. ProDrop utilizes the voxel method to represent the 3D density distribution and the ray tracing approach described in Section 2.1 to find the various projections. The program is highly portable and has been optimized to produce 300 by 300 pixel images in 30 seconds on a Cray II supercomputer. We are currently investigating the possible appearance of gyroid-like morphologies in certain block copolymer samples.

5 Conclusion

We have presented three methods for computing simulations of TEM projections, and a method for obtaining complete registration between the simulated and experimental images. All the methods can handle complicated structures that are determined by highly-connected dividing surfaces. One method uses an adaptation of a ray-tracing algorithm to compute model projections. It divides each projection ray into a series of segments by intersecting it with the dividing surface, which is represented as a mesh of triangles. The segments are divided into two classes corresponding to the two phases in the model, using the pairity of the triangle-ray intersections. The proportion of the lengths of the two classes of segments determines the color of the pixel corresponding to the ray. An oct-tree subdivision scheme is employed to access the triangles and determine any ray-triangle intersections in an efficient way. Another method is relatively simple and can easily be adapted to a wide variety of representations of the dividing surface. In particular, we have demonstrated projection ray subdivision method in one case of a finite element representation that illustrates the versatility in finite element parametrizations which was elaborated by Kistler and Scriven [8].

The results can be dramatic, even in cases where the complexity of the structure makes it extremely difficult for the user to predict the appearance of particular projections. The facility of the method in obtaining a quantitative match is due to two reasons: 1) arbitrary projections can be simulated quickly, thus permitting convenient perusal and cataloging of images; and 2) the simulated annealing relaxation scheme automatically searches for optimal fit parameters. Of course, in many cases visual (i.e., qualitative) matches will be sufficient, in which cases certain of the registration parameters such as the viewing direction (and often the microtome planes) will be important, and others such as the magnification, rotation and shift parameters will not be important. In such cases the 14-dimensional search space may reduce to a much smaller space.

The method can easily be adapted to, for example, structures defined by three-dimensional density distributions, or by dividing surfaces which are the mid-surfaces of continuous density gradients. If the density can be represented as $\rho = \rho(x, y, z)$ or in some other three-dimensional (not necessarily orthogonal) coordinate representation, then this would be explicitly entered in place of the tests described herein. Or, if the density were given as a function of the distance from the dividing surface, then smoothly-varying density gradients, such as those in weakly-segregated systems, could be treated.

Acknowledgements. Financial support from the University of Massachusetts Institute for Interface Science (supported by IBM), the National Science Foundation, grants DMR-89-07433 (ELT) and DMS-8810088 (DH) the Center for Materials Science and Engineering at MIT, and the Department of Energy grant DE-FG02-86ER25015 is gratefully acknowledged.

References

[1] R. C. Gonzales and P. Wintz. *Digital Image Processing.* Addison-Wesley, Reading, MA, 1987.

[2] R. W. Gerchberg. *Opt. Acta*, 21:709, 1974.

[3] A. Papoulis. *IEEE Trans. Circuits, Syst.*, page 735, 1975.

[4] J. M. Carazo and J. L. Carrascosa. *J. Microscopy*, 145(2):159, 1986.

[5] J. M. Cowley and A. F. Moodie. *Acta Cryst.*, 10:609, 1957. Arizona State University, MULTISLICE Program.

[6] E. L. Thomas, D. B. Alward, D. J. Kinning, D. C. Martin, D. L. Handlin, and L. J. Fetters. *Macromolecules*, 19:2197, 1986.

[7] J. R. Bellare, H. T. Davis, W. G. Miller, and L. E. Scriven. Polarized optical microscopy of anisotropic media: imaging theory and simulation. *J. Colloid Interface Sci.*, 136(2):305–326, 1990.

[8] S. Kistler and L. E. Scriven. Computational analysis of polymer processing. Applied Science Publishers Ltd., Barking, Essex, England. Chapter 8.

[9] D. M. Anderson, D. C. Martin, and E. L. Thomas. Maximum entropy data restoration using both real and Fourier space analysis. *Acta Crystallographica, A45*, pages 686–698, 1989.

[10] D. M. Anderson, H. T. Davis, J. C. C. Nitsche, and L. E. Scriven. Periodic surfaces of prescribed mean curvature. Advances in Chem. Phys.

[11] D. M. Anderson. *Studies in the Microstructure of Microemulsions.* PhD thesis, University of Minnesota, June 1986.

[12] H. A. Schwarz. *Gesammelte Mathematische Abhandlungen*, volume 1. Springer, Berlin, 1890.

[13] H. S. M. Coxeter. *Regular Polytopes.* MacMillan, New York, 1963.

[14] H. F. Scherk. Bemerkungen über die kleinste Fläche Innerhalb gegebener Grenzen. *J. R. Angew. Math.*, 13:185–208, 1835.

[15] J. C. C. Nitsche. *Lectures on Minimal Surfaces*, volume 1. Cambridge University Press, 1989.

[16] M. Callahan, D. Hoffman, and J. Hoffman. Computer graphics tools for the study of minimal surfaces. *Communications of the ACM*, 31(6):648–661, 1988.

[17] H. G. Von Schnering and R. Nesper. *Angewandte Chemie, International Editions in English*, 11:1059, 1987.

[18] I. S. Barnes, S. T. Hyde, and B. W. Ninham. The cæsium chloride zero potential surface is not the Schwarz P-surface. *Colloque de Physique*, 7:19–24, 1990. International Workshop on Geometry and Interfaces, E. Dubois-Violette and B. Pansu, Aussois, France.