Model the Solvent-Excluded Surface of 3D Protein Molecular Structures Using Geometric PDE-Based Level-Set Method

Qing Pan[1] and Xue-Cheng Tai[2,3,*]

[1] *College of Mathematics and Computer Science, Hunan Normal University, Changsha, 410081, China.*
[2] *Division of Mathematical Sciences, School of Physical and Mathematical Sciences, Nanyang Technological University, Singapore.*
[3] *Department of Mathematics, University of Bergen, Norway.*

Abstract. This paper presents an approach to model the solvent-excluded surface (SES) of 3D protein molecular structures using the geometric PDE-based level-set method. The level-set method embeds the shape of 3D molecular objects as an isosurface or level set corresponding to some isovalue of a scattered dense scalar field, which is saved as a discretely-sampled, rectilinear grid, i.e., a volumetric grid. Our level-set model is described as a class of tri-cubic tensor product B-spline implicit surface with control point values that are the signed distance function. The geometric PDE is evolved in the discrete volume. The geometric PDE we use is the mean curvature specified flow, which coincides with the definition of the SES and is geometrically intrinsic. The technique of speeding up is achieved by use of the narrow band strategy incorporated with a good initial approximate construction for the SES. We get a very desirable approximate surface for the SES.

1 Introduction

The research of proteins brings the surprising information that the geometric shape of proteins is very important in their function. Proteins are considered as 3D objects with the

---

*Corresponding author. *Email addresses:* `panqing@lsec.cc.ac.cn` (Q. Pan), `Tai@mi.uib.no` (X.-C. Tai)

envelope of surface. Most proteins function through interactions with each other at their molecular surfaces (interfaces). Hence, a proper description of the protein molecular surfaces is essential to the understanding of the life process. Moreover, it also has important implications for the calculations in protein folding, stability, docking and structure-based drug design. The geometric details of the molecular surfaces, called active sites, such as pockets, protrudes and cavities etc., can open up a patch of studying the life function from the viewpoint of the molecular structures. Therefore, the molecular modeling requires special concerns in defining and representing the molecular surfaces of proteins on which interactions occur. There are three common definitions of the molecular surfaces: the Van Der Waals (vdW) surface, the solvent-accessible surface (SAS) and the solvent-excluded surface (SES). The SES is preferred over the SAS as a more suitable description of the topology and details on the surface of proteins, as well as the significant implications in protein recognition, energetics, and stability calculations.

There have been a number of techniques in molecular structure modeling. Unstructured mesh based on the Voronoi-Delaunay framework and the advancing front technique was presented in [30]. Molecular skin model, an explicit triangular description with shape and topology adaptation, was introduced in [18, 19]. A spline parametric approximation method using solvent probe molecules with variable radius was proposed in [8, 9]. A compressed volumetric representation of biomolecular structures was given in [10]. A novel concept of minimal molecular surfaces is presented in [2]. A level-set method [12] based on the solvation free energy functional in the variational implicit solvent model is used to numerically capture arbitrarily shaped solute-solvent interfaces for biomolecular models in the aqueous environment. A level-set front-propagation method to identify the molecular surface and detect the interior cavities within protein structure by a unified and efficient work is adopted in [11]. Based on the Gaussian density map method [4, 15, 16], molecular surfaces are approximated as an isocontour of implicit solvation models. Recently a mesh generation combining a modified dual contour method with implicit solvation model in the form of the Gaussian density map is implemented in [42], which extracts triangular and interior/exterior tetrahedral meshes of molecular structures. Generating a high-quality and adaptive unstructured mesh for molecular structures is very important and challenging because of their applications in so many areas.

Since the level-set method was invented by Osher and Sethian [28, 29, 36], it has been successfully applied to a broad range of problems, such as interface evolution, fluids, materials, combustion, image processing and computer vision, etc.. The level-set method provides a powerful strategy for deforming surfaces, which can easily handle complicated topologies and topologies changing (split apart or merge together), as well as noisy or highly non-uniform data sets. It treats a 3D surface as the level-set of a 3D scalar function, that is, embeds the deformable surfaces as the level surface of a volume, and an evolution equation on the volume governs the behavior of the level surfaces that lie within it. It has a lot of advantages in their applications in 3D graphics. For surface morphing [21, 23, 32], the level-set method can be applied to a wide range of shapes

and topologies, and easily incorporate global coordinate transformations that are generated automatically or by user input, which are difficult for other existing numerical methods. In the filling and blending technique [5, 14, 27], it can generate surfaces with no requirements about prior knowledge of the shapes and topologies for the involved objects. Moreover, the results are invariant, independent of the choice of coordinate systems. They are geometrically intrinsic. For 3D reconstruction problems with no unique solution [1, 13, 39], the level-set method offers a mechanism for finding the optimal solution surface with the invariant smoothness as well as a broad flexibility in shaping the topologies. The drawback of the level-set method, the computational expense, as well as its advantages, have been well known. The approach embeds the interface as the level-set of a higher dimensional function. For two dimensional curves, the computational complexity is $\mathcal{O}(N^2)$ (N stands for the size of the grid in one dimensional direction), and arrives at $\mathcal{O}(N^3)$ in the case of three dimensional surfaces. To speed up the level-set methods, particularly for higher dimensional problems, a localization technique was developed in [20, 31, 35]. The basic idea is to operate only on a neighboring band around the region of interest, that is, around the level-set being tracked (e.g., the level set at a certain value). It can considerably reduce the complexity to $\mathcal{O}(kN)$ in the 2D case and $\mathcal{O}(kN^2)$ in the 3D case, where k stands for the width of the narrow band, and N is the size of the grid in one dimensional direction. There are plenty of research on using geometric partial differential equations(GPDEs) to handle the surface modeling problems in [6, 17, 24, 26, 38, 40, 41]. PDEs are solved on an implicit surface represented by the level-set function in [3, 7].

Our work is modeling the SES of 3D protein molecular structures by use of the geometric PDE-based local level-set method. The GPDE we use is the mean curvature specified flow which coincides with the definition of the SES, moreover, it is geometrically intrinsic. We adopt c^2 smooth tri-cubic spline functions as the level-set function, which is reasonable and improves the quality of the modeled surfaces. The obtained surfaces are very desirable, which can improve the efficiency and quality of capturing the geometric feature of the SES, such as pockets, cavities and protrudes etc.. We also achieve a significant speeding up by use of the narrow band technique combined with the good initial approximate surface constructed from the Gaussian density map method.

The rest of the paper is organized as follows: Section 2 describes the complete algorithm steps in this paper, including the GPDE and the reinitialization equation. Section 3 gives the numerical implementations. In Section 4, we represent the results of our method. Section 5 is for discussion and future work.

2 Algorithm steps

In the life world bio-molecules lie in the solvent (typically water). An interface (surface) separates the bio-molecules from their surrounding solvent molecules because of the chemical interactions between them. Researchers have given three common defini-

tions to describe the interfaces. We use Fig. 1 (a) to explain the three types of the interfaces. Every atom is modeled as a rigid sphere with its Van Der Waals radius. A molecule is an object in the space accumulated by its atoms which naturally generate a surface for the molecule. The simplest is the Van Der Waals (vdW) surface, which is merely the collection of accumulated atomic sphere surface. The solvent-accessible surface (SAS) is the trace left by the center of a solvent probe molecule, also modeled as a rigid sphere when it rolls around the vdW surface, and essentially an extended vdW surface with the solvent probe molecular radius. The solvent-excluded surface (SES) which is composed of the convex and the concave patches, is the boundary of the region inaccessible to a solvent probe sphere as it rolls over the molecule. According to the description of the SES, it is composed of two distinct parts. One part is the contact surface (convex surface), which is the part of the vdW surface of the molecule that is in contact with the solvent probe molecules, and is C^∞ smooth. The other is the reentrant surface (concave surface), which consists of the interior facing portions of the solvent probe sphere that touches two or more atoms simultaneously, and is C^1 or C^0 continuous. It is important to note that the SES is defined by the atoms of protein molecules and the types of solvent probe molecules.
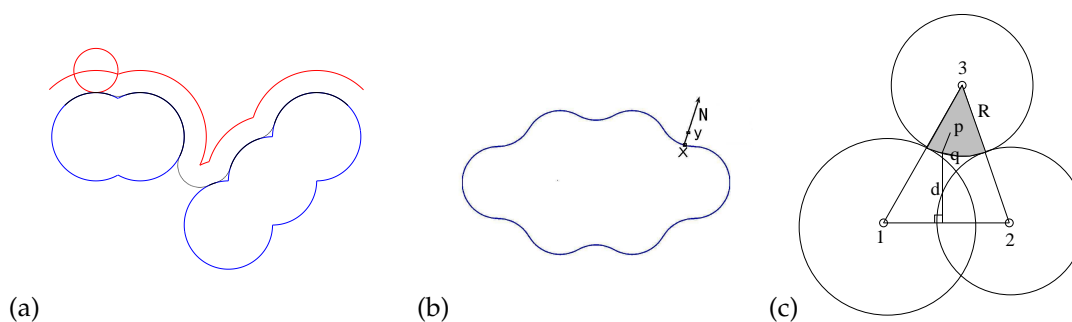

Figure 1: (a). The blue curve depicts the vdW surface of the molecular structure. The red circle represents the solvent probe molecule, and the red curve depicts the SAS. The black curve depicts the SES. (b). From a point $\mathbf{y}$ out of the SES, we introduce the normal N outward the SES. We specify the mean curvature $H_0(\mathbf{y})$ at $\mathbf{y}$ to be equal to the mean curvature $H_0(\mathbf{x})$ at $\mathbf{x}$ which is the intersection point of the normal N and the SES. (c). The spheres labeled with '1' and '2' are the atomic spheres of the molecular structure. The sphere labeled with '3' is the solvent probe molecular sphere. p is a volumetric grid point lying in the shaded domain, which is generated when the solvent probe molecular sphere simultaneously touches two atoms and the two atoms intersect each other, and q is its nearest point on the solvent probe molecular sphere. d is the vertical distance from q to the line connecting the centers of atomic sphere '1' and '2', and R is the radius of the solvent probe molecular sphere.

The topology structure of the SES is very complicated as we see from the above description. It is very difficult to track the topology change of the SES if we use the parametric form PDEs, however, implicit form PDEs can easily overcome this difficulty. We dynamically deform and track the level-set surface driven by our implicit form PDE, which describes the geometric character of the SES. In this paper, we consider using C^2 tri-cubic spline functions as the level-set function, not using C^1 tri-linear spline functions which is obviously easier to implement. There are some reasons for this. Firstly, our

GPDE is driven by the mean curvature, which needs accurate estimate for the derivatives and the mean curvature on the surface, so C^2 smooth cubic spline functions are desirable. It is not reasonable to use higher order splines because the tri-cubic spline is the lowest order spline with C^2 continuity. Secondly, the target surface has certain smoothness requirement except on some finite domains. Using the C^2 smooth level-set function could automatically satisfy this demand, at the same time, we can better treat some singularity which have bad level-set values. Thirdly, it is much more economic from the point of view of the computation costs using the tri-cubic spline functions than increasing the volumetric grid resolution.

Assume that $\mathcal{M}(t)$ is a closed moving interface in IR^3, which encloses a region $\Omega(t)$. We describe $\Omega(t)$ with an auxiliary function $\Phi(\mathbf{x},t)$, called the (scalar) level set function, where $\mathbf{x} \in IR^3, t > 0$. Let Φ move with an appropriate velocity. If we know Φ at any time, we may locate the interface $\mathcal{M}$ by solving the value of the level set function Φ, that is,

$$\mathcal{M}(t) = \{\mathbf{x} \in IR^3 : \Phi(\mathbf{x},t) = c\}.$$

We represent the SES as the zero isocontour of the level set function $\Phi(\mathbf{x},t)$ when $t \to \infty$. When Φ arrives at its steady state, we can capture the SES by solving $\Phi(\mathbf{x},t) = 0$.

Given a uniform volumetric control grid $\{V_{i,j,k}\}_{i,j,k=1}^{L,M,N} := \{x_i\}_{i=1}^{L} \times \{y_j\}_{j=1}^{M} \times \{z_k\}_{k=1}^{N}$, we use tri-cubic tensor product B-spline to approximate the level-set function Φ, i.e.

$$\Phi(\mathbf{x},t) = \Phi(u,v,w,t)$$
$$= \sum_{i=1}^{L}\sum_{j=1}^{M}\sum_{k=1}^{N} \beta_{i,3}(u)\beta_{j,3}(v)\beta_{k,3}(w)d_{i,j,k}(t), \quad u,v,w \in [0,1], \tag{2.1}$$

where $\beta_{i,3}(u)(i=1,\cdots,L)$, $\beta_{j,3}(v)(j=1,\cdots,M)$ and $\beta_{k,3}(w)(k=1,\cdots,N)$ are the class of the third-order B-spline basis functions respectively in x-, y- and z-directions of the volume, and $d_{i,j,k}$ is the values to be determined at volumetric grid points $V_{i,j,k}(i=1,\cdots,L;j=1,\cdots,M;k=1,\cdots,N)$.

The outline of our main algorithm is as follows:

Algorithm 2.1:

---

1. Obtain the initial level-set function values $\Phi(\mathbf{x},0)$, then re-initialize it to be a signed distance function.

2. Update the level-set function $\Phi(\mathbf{x},t)$ using the governing GPDE (2.4).

3. Apply the re-initialization step to set the level-set function $\Phi(\mathbf{x},t)$ to be the signed distance function.

4. Repeat step 2 and 3 until the stopping criterion (3.5) is satisfied.

5. Extract the triangular mesh of the iso-surface $\Phi(\mathbf{x},t) = 0$.

In the following sections, we will describe these steps in details.

2.1 Evolution equation

After we obtain the initial level-set function, it will be followed by the evolution of the level-set function. We now describe the geometric PDE model used in this paper. Let $\mathcal{M}(0)$ be a compact closed immersed orientable surface in IR^3. The curvature driven geometric evolution consists of finding a family $\{\mathcal{M}(t) \geq 0 : t \geq 0\}$ of smooth closed immersed orientable surface in IR^3 according to the flow equation

$$\frac{\partial \mathbf{x}(t)}{\partial t} = V_n(k_1, k_2, \mathbf{x}) N(\mathbf{x}), \tag{2.2}$$

where $\mathbf{x}(t)$ is a surface point on $\mathcal{M}(t)$, $V_n(k_1, k_2, \mathbf{x})$ denotes the normal velocity of $\mathcal{M}(t)$, which depends on the principal curvatures k_1, k_2 of $\mathcal{M}(t)$, and $N(\mathbf{x})$ stands for the unit normal of the surface at $\mathbf{x}(t)$. If taking the normal velocity $V_n(k_1, k_2, \mathbf{x}) = H - H_0(\mathbf{x})$, we get the mean curvature specified flow

$$\frac{\partial \mathbf{x}(t)}{\partial t} = [H - H_0(\mathbf{x})] N(\mathbf{x}), \tag{2.3}$$

where $H_0(\mathbf{x})$ is the specified mean curvature scalar field describing the geometric properties of the SES.

If we use the level set function $\Phi(\mathbf{x}, t)$ to represent the surface, the mean curvature specified flow in the level-set form is

$$\frac{\partial \Phi(\mathbf{x}, t)}{\partial t} = \left[\text{div} \left(\frac{\nabla \Phi(\mathbf{x}, t)}{|\nabla \Phi(\mathbf{x}, t)|} \right) - 2H_0(\mathbf{x}) \right] |\nabla \Phi(\mathbf{x}, t)|. \tag{2.4}$$

The above formula arrives at its steady state when $t \to \infty$, and the isocontour of the level set function Φ must satisfy the condition that the first right-hand term of (2.4) is equal to zero.

Now we describe the mean curvature scalar field on the SES. For a point on the contact surface, its mean curvature is the reciprocal of the ith atomic radius it lies on. For a point on the reentrant surface, its mean curvature is the reciprocal of the solvent probe molecular radius. There is a special case that for a point on the reentrant surface when the solvent probe molecular sphere simultaneously touches two atoms and the two atoms intersect each other(see Fig. 1(c)), its mean curvature is set to be the half of difference between the reciprocal of the vertical distance from this point to the line connecting the two atomic centers and the reciprocal of the solvent probe molecular radius. For a point on the reentrant surface when the solvent probe molecular sphere simultaneously touches more than two atoms and these atoms intersect each other, its mean curvature is the reciprocal of the solvent probe molecular radius. Moreover, we need specify the mean curvature $H_0(\mathbf{y})$ at some point $\mathbf{y}$ out of the SES. We introduce the normal from this point toward the SES and then get the intersection point $\mathbf{x}$ on the SES. We set the mean curvature $H_0(\mathbf{y})$ at point $\mathbf{y}$ to be equal to the mean curvature $H_0(\mathbf{x})$ at point $\mathbf{x}$ (see Fig. 1 (b)).

In general, Φ will deviate away from a signed distance function after several steps of evolution. It needs a reinitialization procedure to keep the character of the signed distance function $|\nabla\Phi| = 1$ and does not change the interface motion. A simple initial value problem [31, 33, 37]

$$
\begin{cases}
\dfrac{\partial \Phi}{\partial \eta} = S(\Phi_0)(1 - |\nabla\Phi|), \\
\Phi(\mathbf{x}, 0) = \Phi_0
\end{cases}
\tag{2.5}
$$

can satisfy this requirement. For the convenience of computation, the sign function $S(\Phi_0)$ is smoothed as

$$
S(\Phi_0) = \frac{\Phi_0}{\sqrt{\Phi_0^2 + \epsilon^2}},
$$

where ϵ is a very small value and chosen to be $\mathcal{O}(\triangle x)$.

3 Numerical implementation

3.1 Initial construction

The data of the complex 3D protein molecular structures, saved as pdb format files, can be downloaded from the free website of the protein data bank: `http://www.rcsb.org`. In this paper, we only need the coordinates of all the major atoms of the molecular structures, whose hydrogen atoms are not given in the pdb files, and their radii which can be known according to the types of the atoms.

In our numerical experiment, a volume which is big enough to embed a molecule is given as

$$
[-D_x, D_x] \times [-D_y, D_y] \times [-D_z, D_z]
$$

with the uniform grid size $L \times M \times N$. The set of volumetric grid points

$$
\{V_{i,j,k}\}_{i,j,k=1}^{L,M,N} := \{x_i\}_{i=1}^{L} \times \{y_j\}_{j=1}^{M} \times \{z_k\}_{k=1}^{N}.
$$

In general, the interface $\mathcal{M}(t) = \{\mathbf{x} : \Phi(\mathbf{x},t) = 0\}$ does not depend on the particular choice of the initial value $\Phi(\mathbf{x},0)$, and we do not have to know any prior information for the topology of the deformed surfaces. However, a good initial approximation is important for the efficiency of our GPDE. We obtain the initial values d^0 at all volumetric grid points with the Gaussian density map method

$$
G(\mathbf{x}, C) = \sum_{i=1}^{N} e^{C(|\mathbf{x} - \mathbf{x}_i|^2 - r_i^2)} - 1,
\tag{3.1}
$$

where $(\mathbf{x}_i, r_i)$ are the center and the radius of the ith atom of a molecular structure with N atoms. The initial molecular surface is approximated by the isocontour $M(0) = \{\mathbf{x} \in IR^3 : G(\mathbf{x}, C) = 0\}$. This algorithm can quickly generate a good smooth envelope of surface for

the molecular objects. It should be noted that the constant C must be given to be negative so that the first term of (3.1) goes to zero as $|\mathbf{x} - \mathbf{x}_i|$ becomes infinite. We will show the inflation effect of the constant C in Section 4. The conclusion is that the larger C will lead to the greater inflation effect of the envelope of surface for the molecular objects.

3.2 Numerical solution of geometric PDE

After we obtain the good initial approximate surface through the Gaussian density map method, we use the geometric PDE-based level-set method as a numerical tool to deform the implicit surface. Before the evolution procedure, we need to precompute the mean curvature H_0 for every volumetric grid point $V_{i,j,k}$ ($i = 1, \cdots, L$; $j = 1, \cdots, M$; $k = 1, \cdots, N$), which has its individual fixed value and its sign depending on the orientation of the surface. We define the direction of the outward normal of the surface as its positive direction. there are three types of specification manners for the mean curvature H_0 on the volumetric grid points. For a volumetric grid point which is nearest to the ith atomic sphere of the molecular structure, we set its mean curvature H_0 to be $1/r_i$ where r_i is the radius of this atomic sphere, whose sign agrees with the orientation of the surface. For a volumetric grid point which is nearest to the solvent probe molecular sphere, we set its mean curvature H_0 to be $-1/R$ where R is the radius of the solvent probe molecular sphere, whose sign is opposite to the orientation of the surface. For a volumetric grid point p lying in the shaded domain (see Fig. 1 (c)) resulted from the case that the solvent probe molecular sphere simultaneously touches two atoms and the two atoms intersect each other, we find its nearest point q on the reentrant surface, and d is the vertical distance from q to the line connecting the centers of the two atomic spheres, whose sign agrees with the orientation of the surface. The mean curvature H_0 at point p is set to be $\frac{1}{2}(d^{-1} - R^{-1})$. Moreover, for a volumetric grid point lying in the similar domain to the third type resulted from the case that the solvent probe molecular spheres simultaneously touches more than two atoms and the atoms intersect each other, we set its mean curvature to be $-1/R$.

Now we start the deformation procedure governed by Eq. (2.4). For simplicity, we use $\mathcal{I}$ ($1 \leq \mathcal{I} < L \times M \times N$) to denote the index of some point $V_{\mathcal{I}}$ on the set of the volumetric grid points $\{V_{i,j,k}\}_{i,j,k=1}^{L,M,N}$, then rewrite the formula (2.1) as

$$\Phi_{\mathcal{I}} = \sum_{l=1}^{L \times M \times N} \mathcal{B}_l(u,v,w) d_l, \tag{3.2}$$

where $\Phi_{\mathcal{I}}$ is the level-set function value on the $\mathcal{I}$th volumetric grid point, and $\mathcal{B}_l(u,v,w)$ is the corresponding tri-cubic B-spline tensor product $\beta_{i,3}(u)\beta_{j,3}(v)\beta_{k,3}(w)$.

Given a time steplength $\tau > 0$, suppose we have the approximate values d^n of all volumetric grid points at $t_n = n\tau$. We will construct the approximate solution for the next step $t_{n+1} = t_n + \tau$ using the semi-implicit Euler scheme of (2.4)

$$\frac{\Phi^{n+1} - \Phi^n}{\tau} = \left[\text{div} \left(\frac{\nabla \Phi^{n+1}}{|\nabla \Phi^n|} \right) - 2H_0 \right] |\nabla \Phi^n|.$$

Consequently,

$$\Phi^{n+1} - \tau \mathrm{div}(\nabla \Phi^{n+1}) = \Phi^n - 2\tau H_0 |\nabla \Phi^n|. \tag{3.3}$$

Let $D^{n+1} = [d_1^{n+1}, \cdots, d_m^{n+1}]^T \in IR^m$ where $d_1^{n+1}, \cdots, d_m^{n+1}$ are the unknowns to be determined and we use J denote the index set of them. Substitute (3.2) into the above formula, then get the matrix form

$$MD^{n+1} = B^n - C^n, \tag{3.4}$$

where $M = (m_{ij}) \in IR^{m \times m}$, $B^n = (b_i^n) \in IR^m$ and $C^n = (c_i^n) \in IR^m$. We write the elements of the matrix M, B^n and C^n as follows:

$$m_{ij} = \sum_{j \in J} \mathcal{B}_j(u,v,w) - \tau \mathrm{div}\Big[\nabla \sum_{j \in J} \mathcal{B}_j(u,v,w) \Big],$$

$$b_i^n = \Phi_{\mathcal{I}}^n - 2\tau H_0(V_{\mathcal{I}}) |\nabla \Phi_{\mathcal{I}}^n|,$$

$$c_i^n = \sum_{j \notin J} \mathcal{B}_j(u,v,w) d_j^n - \tau \mathrm{div}\Big[\nabla \sum_{j \notin J} \mathcal{B}_j(u,v,w) d_j^n \Big],$$

where $\mathcal{I}$ is the index of the ith unknown value to be determined in the set J, and c_i^n comes from the known terms of the left-handed side of (3.4). The coefficient matrix M is highly sparse, hence a robust iterative method for solving it is desirable. We use Saad's iterative method [34], named GMRES, to solve the linear system (3.4). The equation requires the time steplength $\tau = \mathcal{O}(\triangle x^2)$. We stop the equation evolution when the following condition is satisfied

$$\|\mathcal{M}(t_{n+1}) - \mathcal{M}(t_n)\| / \tau \le \epsilon_0, \tag{3.5}$$

where ϵ_0 is a given small constant.

3.3 Numerical solution of reinitialization

There are a lot of numerical methods for the reinitialization equation (2.5). We adopt the modified Godunov discrete formula [22, 33].

Suppose we have the approximate solution $\Phi_{i,j,k}^n$ of the system (3.3) at the time $t = n\tau$. We replace $\Phi_{i,j,k}^n$ with $\Phi_{i,j,k}^{\nu,n}$ ($\nu \ge 0$), then get the discrete formula of (2.5)

$$\Phi_{i,j,k}^{\nu+1,n} = \Phi_{i,j,k}^{\nu,n} - \eta S(\Phi) G(\Phi), \quad \nu \ge 0, \tag{3.6}$$

where

$$S(\Phi) = \frac{\Phi_{i,j,k}^{\nu,n}}{\sqrt{(\Phi_{i,j,k}^{\nu,n})^2 + |\nabla \Phi_{i,j,k}^{\nu,n}|^2 \triangle x^2}},$$

$$G(\Phi) = \begin{cases} \sqrt{\max((a^+)^2,(d^-)^2) + \max((b^+)^2,(e^-)^2) + \max((c^+)^2,(f^-)^2)} - 1, & \Phi_{i,j,k}^{\nu,n} > 0, \\ \sqrt{\max((a^-)^2,(d^+)^2) + \max((b^-)^2,(e^+)^2) + \max((c^-)^2,(f^+)^2)} - 1, & \Phi_{i,j,k}^{\nu,n} < 0, \\ 0, & \text{else,} \end{cases}$$

with

$$a = \frac{\Phi_{i,j,k}^{v,n} - \Phi_{i-2,j,k}^{v,n}}{2\triangle x}, \quad b = \frac{\Phi_{i,j,k}^{v,n} - \Phi_{i,j-2,k}^{v,n}}{2\triangle y}, \quad c = \frac{\Phi_{i,j,k}^{v,n} - \Phi_{i,j,k-2}^{v,n}}{2\triangle z},$$

$$d = \frac{\Phi_{i,j,k}^{v,n} - \Phi_{i+2,j,k}^{v,n}}{2\triangle x}, \quad e = \frac{\Phi_{i,j,k}^{v,n} - \Phi_{i,j+2,k}^{v,n}}{2\triangle y}, \quad f = \frac{\Phi_{i,j,k}^{v,n} - \Phi_{i,j,k+2}^{v,n}}{2\triangle z},$$

and $p^+ = \max(p,0)$, $p^- = \min(p,0)$.

We operate enough many iteration steps to satisfy the numerical requirement

$$\sum_{i,j,k} |\Phi_{i,j,k}^{v+1,n} - \Phi_{i,j,k}^{v,n}| < \epsilon, \tag{3.7}$$

where ϵ is chosen to be $0.01\triangle x$.

3.4 Iso-surface extraction

The surface $\mathcal{M}(t)$ is represented as level-sets of a higher-dimensional tri-cubic spline surface. The initial level-set surface is constructed through the Gaussian density map method. A family of level surfaces is obtained by evolving the level-set function $\Phi(\mathbf{x},t)$ which obeys the rule of the mean curvature specified flow, where t parameterizes the family and $\mathbf{x}$ parameterizes the surface. Therefore, it is easy to extract the desirable SES as an iso-surface $\Phi(\mathbf{x},t) = 0$.

There are a lot of research achievements about the algorithm of the fast and efficient iso-surface extraction. In this paper we adopted the marching cube method [25] to extract the iso-surface of our level-set function.

3.5 Localization

The fast localization technique makes the three dimensional problems not difficult. In this paper, we also implement a local level-set framework to save the computational cost. The domain we need to solve the level set function is a narrow band neighboring the interface $|\Phi| = 0$. To perform it, we only need to update the points whose $|\Phi|$ is less than or equal to a certain ϵ when we solve the evolution equation (2.4), and maintain the level-set function Φ as the signed distance function. We consider the volumetric grid points with the constraint $|\Phi| < k\triangle x$ (k is a integer) as the unknowns to be determined. It tremendously simplifies the computation of the narrow band. Therefore, our algorithm demands a relatively small amount of computational resources.

It should be noted that because the initial level-set function values $\Phi(\mathbf{x},0)$ generated from the Gaussian density map method is far away from the signed distance function, we must apply enough many reinitialization steps to make it to be the signed distance function.
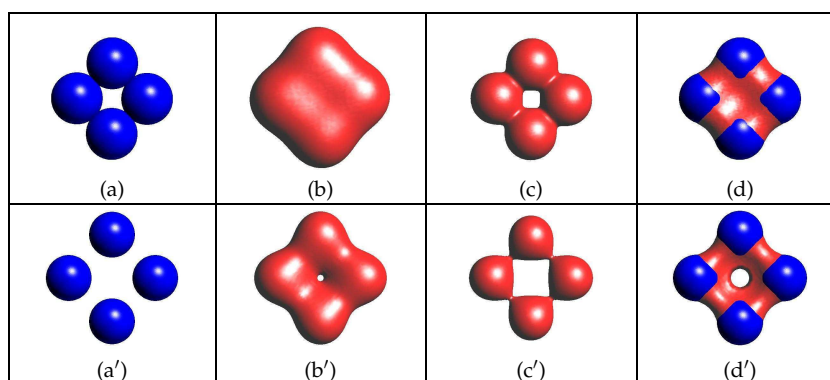
Figure 2: In the first row: (a) shows four same atomic spheres uniformly intersecting each other. (b) and (c) show the envelope of surface for the object constructed using the Gaussian density map method, where we choose $C = -0.3$ in (b) and $C = -1.4$ in (c). (c) shows the more tightly enclosing effect than (b). The envelope of surface is very close to the vdW surface of the object when we choose a very small constant C in (c). (d) is the SES of the object generated from our method, where the middle vacant space among these four atomic spheres disappears, enclosed by the reentrant surface. In the second row: (a′) shows four same atomic spheres uniformly isolated from each other. (b′) and (c′) show the envelope of surface for the object constructed using the Gaussian density map method, where we choose $C = -0.21$ in (b′) and $C = -0.34$ in (c′). (b′) and (c′) show the same result of the enclosing effect for the constant C as (b) and (c). (d′) is the SES of the object generated from our method, where a smooth contour of the water molecular sphere in the middle vacancy among these four atomic spheres appears, and these four atoms are connected through the smooth reentrant surface. We use blue to render the contact surface and red to render the reentrant surface in (d) and (d′). The solvent probe molecular radius is 1.4.

4   Results

In this section, we represent some examples to illustrate the effect of our method.

Fig. 2 shows two examples of four atomic spheres with the same size uniformly intersecting each other and isolated from each other. We represent the envelope of surface with the different enclosing effect for these objects by choosing the different constant C of the Gaussian density map method in (b), (c), (b′) and (c′), and their SES generated from our method in (d) and (d′). It is clear to see that the envelope of surface constructed with the Gaussian density map method is obviously different from the SES.

Fig. 3 shows a simple real protein molecule 6RXN with 369 atoms. We use different colors to render the different residues of the molecule in Fig. 3(a). By use of the Gaussian density map method, we represent three pieces of the envelop of surface for this molecule with the different inflation effect in (b), (c) and (d). These three figures show the effect that the envelope of surface more tightly encloses the molecule with the smaller value of the constant C, as described in Section 3.1. We can see that the envelope of surface of the molecule shown in (d) is very close to its vdW surface when we choose a very small constant C. (e) shows the SES of the molecule generated from our method. (f), where we use the same color system as (a) to render the part of the SES belonging to the vdW surface of atomic spheres, i.e., contact surface, and red to render the rest, i.e., the reentrant surface, shows the same SES as (e).
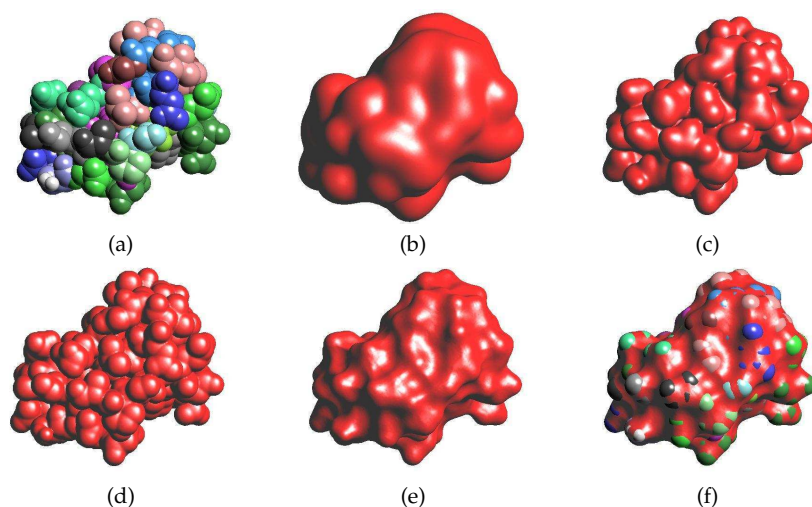
Figure 3: (a) shows a small protein molecule named 6RXN with 369 atoms, whose residues are rendered using different colors. (b), (c) and (d) show the envelope of surface of the molecule resulted from the Gaussian density map method, where we choose $C = -0.2$ in (b), $C = -0.8$ in (c) and $C = -4.0$ in (d). It is clear to see that the inflation effect reduces when the value of the constant C decreases. Finally, the envelope of surface of this molecule is very approximate to its vdW surface when we choose a very small constant C, as shown in (d). (e) and (f) show the same result, the SES generated from our method. In (f) we use the same color system as (a) to render the contact surface of the SES, and red to render the reentrant surface of the SES. The solvent probe molecular radius is 1.4.

Fig. 4 represents two examples of complex protein molecules. The volumetric grid size is chosen to be less than one-sixth of the minimal atomic radius so that the grid resolution guarantees that our level-set function can track every atom. In this figure, we also represent the envelope of surface of these molecules constructed from the Gaussian density map method and the SES produced by use of our method. It should be noted that the envelope of surface generated from the Gaussian density map method is far away from the real SES no matter how you choose the constant C. However, an appropriate

Table 1: Area resulted from the MSMS software and our method.

PDB	Atoms	Msms	Gpde	PDB	Atoms	Msms	Gpde
6RXN	369	2279	2263	1AA2	887	4816	4768
1UCH	1660	8151	8056	2TEC	2526	10294	10170
1HGA	4384	21818	21527	1MAH	4576	18762	18518
1FSS	4689	18724	18481	2C00	6748	29769	29334
2J9G	6865	31572	31066	3ECD	11605	50066	49014

We show ten examples in this table. The first and the fifth columns are the PDB names for the ten molecules. The second and the sixth columns show the number of their atoms correspondingly. The third and the seventh columns show the area resulted from the MSMS software. The forth and the eighth columns show the area resulted from our method. The solvent probe molecular radius is chosen to be 1.4.
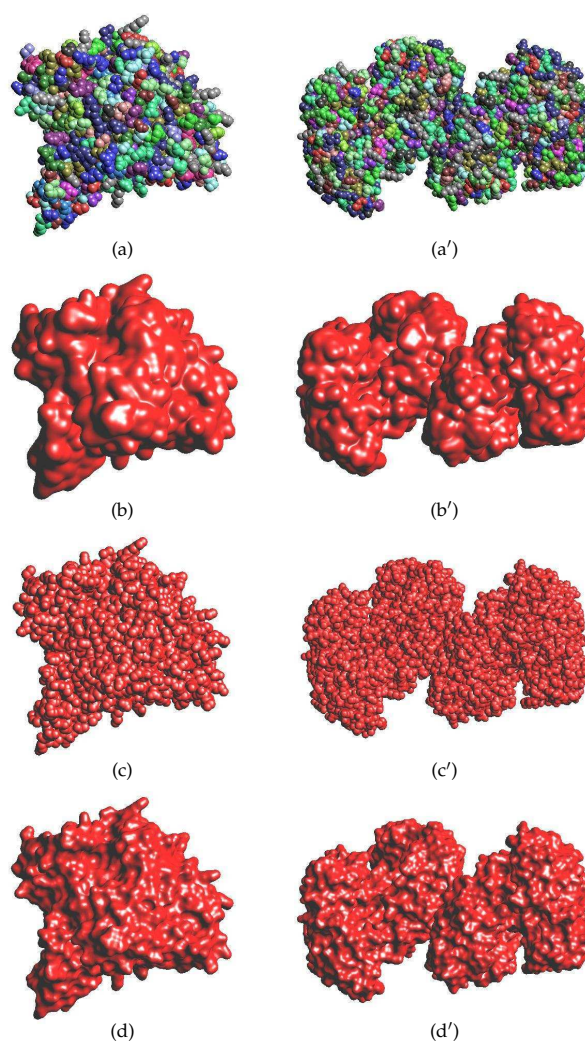
Figure 4: (a) and (a$'$) are the real protein molecules named 1FSS and 3ECD with 4689 and 11605 atoms respectively, where we use different colors to render their different residues. (b) and (c), (c$'$) and (d$'$) are their corresponding envelope of surface resulted from the Gaussian density map method, where we choose $C = -0.2$ in (b) and (b$'$), and $C = -4.0$ in (c) and (c$'$). You can clearly see the contour of their skin atoms in the third row. This effect is in contrast with the bigger inflation effect shown in the second row. (d) and (d$'$) are their corresponding SES generated from our method. The solvent probe molecular radius is 1.4.

constant C can provide a good initial value for the following equation evolution and greatly save the computational cost.

In order to further highlight the advantages of our method, we compare our method with a software called MSMS on ten examples in Table 1. MSMS is a commonly used method for this kind of computations. We compute the sum of area of all extracted triangular patches of the SES generated from our method, and compare it with the data from the MSMS software. It shows that our method is reliable.

5   Discussion and future work

The three-dimensional protein molecules have their individual steric structures, which are always shaped with their geometric details on the skin of them, and the information that the molecular structure finally determines the life phenomena tells us the importance of the accurate and appropriate description for these geometric features. The SES is a general model characterizing the structural shapes, which expresses to some degree the physicochemical properties of molecules, and is broadly applied in many areas. Although the approximate surface resulted from the Gaussian density map method is obviously different from the SES, it provides a good initial approximation construction for our GPDE evolution. In this paper, the geometric PDE-based local level-set method we use can precisely model the SES. The geometric PDE is the mean curvature specified flow, which is derived from the geometric character of the SES. We obtain desirable results. Further improvement of the efficiency of our method will be under our consideration in the future.

There are broader applications of the proposed method after we finish this work. For example, we can use the proposed method to improve the effect of the protein comparison and the classification technique through characterizing the geometric details on the SES, and the accuracy in predicting and recognizing the docking area of protein-protein interactions, which is helpful to probe how proteins interact with each other from the viewpoint of the geometric shape complementarity.

Acknowledgments

References

[1] C. Bajaj, F. Bernardini, and G. Xu. Automatic reconstruction of surfaces and scalar fields from 3d scans. In SIGGRAPH'95 Proceedings, vol. 19(3), pp. 109–118, 1995.

[2] P. Bates, G.W. Wei, and S. Zhao. Minimal molecular surfaces and their applications. J. Comput. Chem., 29:380–391, 2008.

[3] M. Bertalmio, L.-T. Cheng, S. Osher, and G. Sapiro. Variational problems and partial differential equations on implicit surfaces: The framework and examples in image processing and pattern formation. J. Comput. Phys., 174:759–780, 2001.

[4] J. Blinn. A generalization of algebraic surface drawing. ACM Transactions on Graphics, 1(3):235–256, 1982.

[5] J. Bloomenthal and K. Shoemake. Convolution surfaces. In T. W. Sederberg, ed., Computer Graphics (SIGGRAPH'91 Proceedings), vol. 25, pp. 251–257, 1991.

 [6] M. I. G. Bloor and M. J. Wilson. Generating blending surfaces using partial differential equations. Computer Aided Design, 21(3):165–171, 1989.
 [7] P. Burchard, L.-T. Cheng, B. Merriman, and S. Osher. Motion of curves in three spatial dimensions using a level set approach. J. Comput. Phys., 170(2):720–741, 2001.
 [8] C. Bajaj, H. Lee, R. Merkert, and V. Pascucci. NURBS based b-rep models from macromolecules and their properties. In Proceedings of 4th Symposium On Solid Modeling And Applications, pp. 217–228, 1997.
 [9] C. Bajaj, V. Pascucci, A. Shamir, R. Holt, and A. Netravali. Dynamic maintenance and visualization of molecular surfaces. Discrete Applied Mathematics, 127:23–51, 2003.
[10] C. Bajaj and V. Siddavanahalli. Fast and adaptive surfaces and derivatives computations for volumetric particle data. CS and ICES Technical Report, the University of Texes at Austin, 2005.
[11] T. Can, C.-I. Chen, and Y.-F. Wang. Efficient molecular surface generation using level-set methods. Journal of Molecular Graphics and Modelling, 25:442–454, 2006.
[12] L.-T. Cheng, J. Dzubiella, J. A. McCammon, and B. Li. Application of the level-set method to the implicit solvation of nonpolar molecular. J. Chem. Phys., 127(8), 2007.
[13] B. Curless and M. Levoy. A volumetric method for building complex models from range images. In Computer Graphics (SIGGRAPH'96 Proceedings), July 1996.
[14] M. Desbrun and M. P. Gascuel. Animating soft substances with implicit surfaces. In SIGGRAPH'95 Proceedings, pp. 287–290, Aug. 1995.
[15] R.-R. Gabdoulline and R.-C. Wade. Analytically defined surfaces to analyze molecular interaction properties. Journal of Molecular Graphics, 14(6):341–353, 1996.
[16] J. Grant and Pickup B. A gaussian description of molecular shape. Journal of Physical Chemistry, 99:3503–3520, 1995.
[17] H. Du and H. Qin. Dynamic pde-based surface design using geometric and physical constraint. Graph models, 67(1):43–71, 2005.
[18] H. Edelsbrunner. Deformable smooth surface design. Discrete Computational Geometry, 21:87–115, 1999.
[19] H.-L. Cheng, T. Dey, H. Edelsbrunner, and J. Sullivan. Dynamic skin triangulation. Discrete Computational Geometry, 25:525–568, 2001.
[20] J. Helmsen, E. G. Puckett, P. Colella, and M. Dorr. Two new methods for simulating photolithography development in 3D. In SPIE Microlithography IX 1996, pp. 253, 1996.
[21] J. Hughes. Scheduled fourier volume morphing. In Computer Graphics (SIGGRAPH'92 Proceedings), vol. 26(2), pp. 43–46, 1992.
[22] B. Merriman, J. K. Bence and S. Osher. Motion of multiple junctions: a level set approach. Journal of Comput. Phys., 112:334, 1994.
[23] A. Lerios, C. Garfinkle, and M. Levoy. Feature-based volume metamorphosis. In SIGGRAPH'95 Proceedings, pp. 449–456, Aug. 1995.
[24] H. Li and X.-C. Tai. Piecewise constant level set method for multiphase motion. Int. J. Numer. Anal. Model., 4(2):291–305, 2007.
[25] W. E. Lorensen and H. E. Cline. Marching cubes: A high resolution 3D surface construction algorithm. In SIGGRAPH'87: Proceedings of the 14th Annual Conference on Computer Graphics and Interactive Techniques,, pp. 163–169, 1987.
[26] M. Desbrun, M. Meyer, P. Schröder, and A.-H. Barr. Implicit fairing of irregular meshes using diffusion and curvature flow. In SIGGRAPH99, pp. 317–324, Los Angeles, USA, 2002.
[27] A. Middleditch and K. Sears. Blend surfaces for set theoretic volume modeling systems. In Computer Graphics (SIGGRAPH'85 Proceedings), vol. 19(3), pp. 161–170, July 1985.

[28] S. Osher and R. P. Fedkiw. Level-set methods: an overview and some recent results. J. Comput. Phys., 169:475–502, 2001.

[29] S. Osher and J. A. Sethian. Fronts propagating with curvature dependent speed: algorithm based on Hamilton-Jacobi formulation. J. Comput. Phys., 79:12–49, 1988.

[30] P. Laug and H. Borouchaki. Molecular surface modeling and meshing. Engineering with Computers, 18:199–210, 2002.

[31] D.-P. Peng and B. Merriman. A PDE-based fast local level set method. J. Comput. Phys., 155(2):410–438, 1999.

[32] J. Rossignac and A. Kaul. AGRELS and BIPs: metamorphosis as a bezier curve in the space of polyhedra. In Computer Graphics Forum (Eurographics'94 Proceedings), vol. 13(3), pp. 179–184, 1994.

[33] S. Chen, B. Merriman, S. Osher, P. and Semereka. A simple level set method for solving stefan problems. J. Comput. Phys., 135:8–29, 1997.

[34] Y. Saad. Iterative Methods for Sparse Linear Systems. 2nd ed., SIAM, Philadelphia, 2000.

[35] J. A. Sethian. A fast marching level set method for monotonically advancing fronts. Proc. Nat. Acad. Sci., 93, 1591.

[36] J. A. Sethian. Volution, implementation, and application of level set and fast marching methods for advancing fronts. J. Comput. Phys., 169:503–555, 2001.

[37] X.-C. Tai and T. F. Chan. A survey on multiple level set methods with applications for identifying piecewise constant functions. Int. J. Numer. Anal. Model., 1(1):25–47, 2004.

[38] U. Clarenz, U. Diewald, and M. Rumpf. Anisotropic geometric diffusion in surface processing. In Proceedings of Viz2000, pp. 397–505, Salt Lake City, Utah, United States, 2000.

[39] R. T. Whitaker. Algorithms for implicit deformable models. In Fifth International Conference on Computer Vision, IEEE, 1995. IEEE Computer Society Press.

[40] G.-L. Xu and Q. Pan. G^1 surface modelling using fourth order geometric flows. Computer Aided Design, 38(4):392–403, 2006.

[41] G.-L. Xu, Q. Pan, and C. Bajaj. Discrete surface modeling using partial differential equations. Computer Aided Geometric Design, 23(2):125–145, 2006.

[42] Y. Zhang, G. Xu, and C. Bajaj. Quality meshing of implicit solvation models of biomolecular structures. The special issue of Computer Aided Geometric Design on Applications of Geometric Modeling in the Life Sciences, 23(6):510–530, 2006.