Representing Thermal Vibrations and Uncertainty in Molecular Surfaces

Chang Ha Lee and Amitabh Varshney

Department of Computer Science and UMIACS, University of Maryland,
College Park, Maryland, USA

ABSTRACT

The previous methods to compute smooth molecular surface assumed that each atom in a molecule has a fixed position without thermal motion or uncertainty. In real world, the position of an atom in a molecule is fuzzy because of its uncertainty in protein structure determination and thermal energy of the atom. In this paper, we propose a method to compute smooth molecular surface for fuzzy atoms. The Gaussian distribution is used for modeling the fuzziness of each atom, and a p-probability sphere is computed for each atom with a certain confidence level. The smooth molecular surface with fuzzy atoms is computed efficiently from extended-radius p-probability spheres. We have implemented a program for visualizing three-dimensional molecular structures including the smooth molecular surface with fuzzy atoms using multi-layered transparent surfaces, where the surface of each layer has a different confidence level and the transparency associated with the confidence level.

Keywords: molecular surface, atom vibrations, thermal motion, temperature factor, 3D visualization, fuzzy geometry

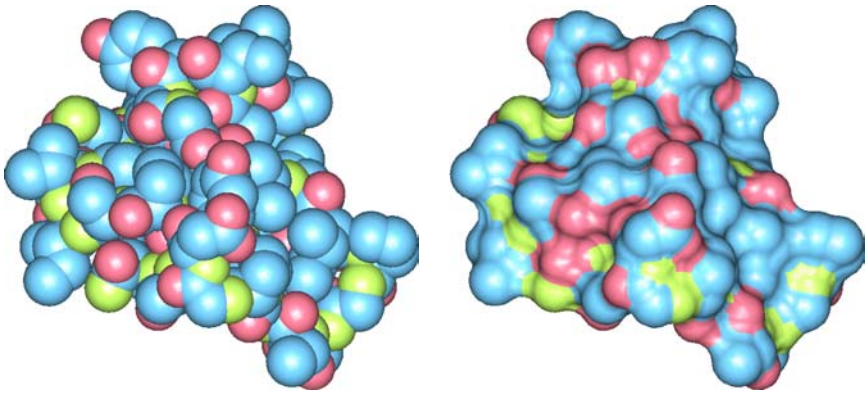
1. INTRODUCTION

The smooth molecular surface, first proposed by Richards,[1] is defined as the surface which an external probe sphere touches when it is rolled over the spherical atoms of a molecule. The smooth molecular surface is useful for studying the structure and interactions of proteins, especially for testing the accessibility of a solvent into a molecule, for prediction of three-dimensional structures of biological macromolecules and assemblies, and for discriminating between different ways of docking molecules which can be used in drug design.

Many researchers have proposed methods to compute the smooth molecular surface (Figure 1), and they have assumed that each atom in a molecule has a certain rigid position. In real world, an atom in a molecule vibrates because of its thermal energy. Therefore, the smooth molecular surface will also vibrate. Also in protein structure determination, sometimes the positions of the atoms are uncertain. This information is useful and should be displayed with the rest of the molecule's visual attributes. In this paper, we propose a method to compute the smooth molecular surface with atoms that are either vibrating or otherwise have uncertainty in their positions. The understanding of thermal motion and uncertainty provides a more informative molecular display and a better understanding of its structure and function.

The remaining paper is organized as follows. In Section 2, we give an overview of the present systems for computing the smooth molecular surface. The fuzzy atom model and the description of computing the isotropic displacement are given in Section 3. We describe how to efficiently compute the fuzzy molecular surface in Section 4, and in Section 5 we discuss the implementation of our algorithm. In Section 6, we conclude this paper and discuss future work.

Further author information: (Send correspondence to C.H.L.)
C.H.L.: E-mail: chlee@cs.umd.edu, Telephone: +1-301-405-1213
Copyright 2001, Society of Photo-Optical Instrumentation Engineers. This paper was published in SPIE Conference on Visualization and Data Analysis, San Jose, CA, January 2002, and is made available as an electronic reprint with permission of SPIE. One print or electronic copy may be made for personal use only. Systematic or multiple reproduction, distribution to multiple locations via electronic or other means, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper are prohibited.



(a) Using spheres to display atoms

(b) Molecular surface (probe radius 1.4)

Figure 1: Molecular visualizations for Crambin

2. PREVIOUS WORK

The earliest algorithms for computing molecular surfaces were proposed by Connolly.[2–4] Each region of the molecular surface is represented by a convex spherical, toroidal, or concave spherical patch when a probe sphere touches one, two, or three atoms respectively. These patches collectively form the molecular surface.

A more efficient molecular surface computation algorithm for interactive visualization is presented by Varshney et al..[5] Their algorithm is analytic and parallelizable. First, a list of neighboring atoms of each atom is determined by finding all the atoms that are close enough to affect the surface for the atom. A feasible cell for each atom is computed by considering its Voronoi neighbors. The patches for each atom i are computed from intersection of the feasible cell for atom i and a sphere whose radius is the radius of atom i added by the probe radius.

Edelsbrunner and Mucke[6] have proposed the concepts of three-dimensional alpha shapes and alpha hulls and have related them to Delaunay triangulations and Voronoi diagrams. Let S be a finite set of points in R^3 and α be a real number with $0 \leq \alpha \leq \infty$. A $k - simplex$, $0 \leq k \leq 3$ is defined as a convex hull of a subset $T \subset S$ of size $|T| = k + 1$. A simplex is an α-exposed simplex if there exists a sphere with radius α touching all of the points in the simplex and not containing any other point in S. The *alpha shape* of a point set S is defined as a polytope whose boundary consists of triangles, edges, and vertices in the set of α-exposed simplices. When α is ∞, the alpha shape is the same as the convex hull of S, and as α decreases, the alpha shape shrinks by gradually developing cavities. Edelsbrunner and Mucke have proposed an efficient algorithm to generate the intervals of simplices that can be used for generating the family of alpha shapes. Molecular surfaces can be modeled by the weighted alpha hulls using their algorithm.

Bajaj et al.[7] have introduced a multiresolution representation scheme for molecular surfaces at various levels of detail. Their scheme provides the flexibility to achieve different space-time efficiency tradeoffs, and tracks the topology of the molecular shape at any adaptive level of resolution. The multiresolution hierarchy is constructed by clustering spheres, each of which represents an atom of the molecule. At each level, spheres are clustered from a priority queue according to the error estimates, and the clustered spheres are replaced with a new bounding sphere at the next level. The center and radius of the new sphere are determined so that the new sphere encloses the clustered spheres. The error estimates used for guiding the clustering process are the Euclidean distance between two center points, the difference in area between the new sphere and the old spheres, and the Hausdorff distance between the boundary of the old spheres and the new one.

Huitema and van Liere[8] have presented the *zonal map*, a data structure for visualizing large, time-dependent molecular conformations. They have defined a focus cell as a cell which is at the center of the user's view. A region of interest with a parameter n is defined as a set of cells within a distance n from the focus cell. To

compute the distance between two cells, they use the L_1 distance metric*. For time-critical computing, the molecular surface for the atoms in the smallest region of interest is initially computed and rendered. The molecular surface for progressively larger regions of interest is then computed and rendered until the target time is reached.

Cai et al.[9] have presented an algorithm for approximating the molecular surface. Their algorithm starts with a triangle mesh built on an ellipsoid embracing the whole molecular surface. The starting ellipsoid is computed using the center of the molecule and its principal directions. The final molecular surface is obtained by iteratively deflating each vertex of the starting triangle mesh until it reaches the molecular surface. The direction of deflation is the average of normal vectors of adjacent triangles, and the amount of deflation is a constant value.

Sanner et al.[10] have described an algorithm that relies on the use of reduced surfaces for computing solvent-excluded (smooth) molecular surfaces. The reduced surface as defined by Sanner et al. corresponds to the alpha shape for that molecule (as defined by Edelsbrunner and Mucke[6]) with a probe radius α. Their algorithm consists of four phases. The first phase computes the reduced surface, the second phase builds an analytical representation of the molecular surface from the reduced surface, the third phase removes all self-intersecting parts from the surface built by the second phase, and the last phase produces a triangulation of the molecular surface.

Klein et al.[11] have introduced a method for generating a molecular surface using parametric patches to represent the surface. First, points are generated on the surface based on a desired distance between points and then approximately equilateral triangles that describe the molecular surface are generated. The control points of the patches which represent the surface are determined from the vertices of the triangles.

3. FUZZY MOLECULES

3.1. Fuzzy Atoms

The position of an atom in a molecule is not stable because of the thermal energy of the atom and uncertainty of the atom position. The PDB (Protein Data Bank) file contains the information of a molecule including the temperature factor for each atom. The probability distribution for each atomic thermal motion can be derived from the temperature factor by computing the mean-square displacement parameters. The atomic thermal motion as well as uncertainty is modeled by a Gaussian distribution:

$$G(X) = (\frac{|U^{-1}|}{(2\pi)^3})^{\frac{1}{2}} e^{-\frac{1}{2}X^T U^{-1} X} \tag{1}$$

In Equation (1), U is the mean-square displacement matrix. For simplicity, for the rest of this paper we shall assume that the thermal motion of an atom is isotropic, i.e. it is the same in all the directions. Thus,

$$G(X) = (\frac{1}{(2\pi)^3} u_{eq})^{\frac{1}{2}} e^{-\frac{1}{2u_{eq}}|X|^2}, \text{ where } U = U_{eq} = \begin{pmatrix} u_{eq} & 0 & 0 \\ 0 & u_{eq} & 0 \\ 0 & 0 & u_{eq} \end{pmatrix} \tag{2}$$

The mean of this distribution is at the origin, $(0, 0, 0)$, the standard deviation is $\sigma = \sqrt{u_{eq}}$, and the amplitude is $\sqrt{\frac{1}{(2\pi)^3} u_{eq}}$.

Because each atom has an arbitrary center and orientation, the general form of the distribution can be defined in the homogeneous coordinate system. Let M be the 4×4 homogeneous 3D transformation matrix for an atom. Also, let $U' = \begin{bmatrix} U & 0 \\ 0 & 1 \end{bmatrix}$. Then, the Gaussian distribution representing the atom with appropriate center is given by

*The L_1 distance metric between two points (x_0, y_0, z_0) and (x_1, y_1, z_1) is given by $|x_1 - x_0| + |y_1 - y_0| + |z_1 - z_0|$

$$G(X) = (\frac{|U'^{-1}|}{(2\pi)^3})^{\frac{1}{2}} e^{-\frac{1}{2}(MX)^T U'^{-1}(MX)}$$
$$= (\frac{|U'^{-1}|}{(2\pi)^3})^{\frac{1}{2}} e^{-\frac{1}{2}X^T(M^T U'^{-1} M)X}$$

Since $U' = U'_{eq}$,

$$G(X) = (\frac{1}{(2\pi)^3} u_{eq})^{\frac{1}{2}} e^{-\frac{1}{2u_{eq}}|MX|^2} \tag{3}$$

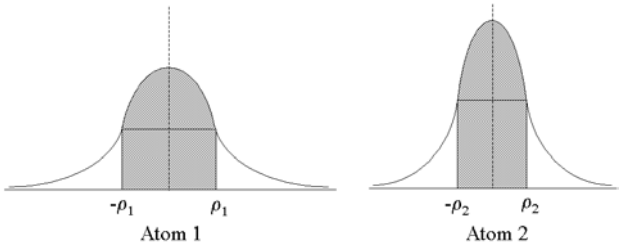
For a fuzzy set u, the level sets[12] (α-cut) of u are given by

$$u_\alpha = \{p \in R^3, u(p) \geq \alpha\}, \quad 0 \leq \alpha \leq 1. \tag{4}$$

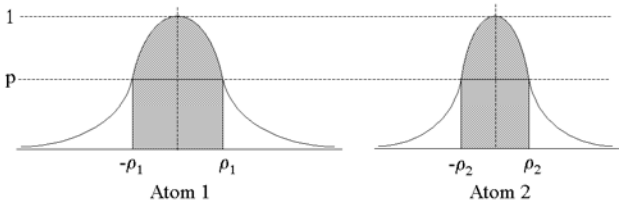
The fuzzy atoms can thus be expressed with a certain confidence level using the level sets of the Gaussian distributions representing the thermal motion of atoms.

3.2. Fuzzy Molecules

Let us consider two atoms whose fuzziness follows Gaussian distributions in figure 2 (a). The shaded area under the Gaussian curves represents the cumulative probability of finding the center of that atom within a distance ρ from its mean center. The shaded regions of the two graphs in Figure 2 (a) represent equal areas and hence equal probabilities of finding the centers of these atoms within distances ρ_1 and ρ_2 from their means, respectively. As it can be seen, since $\rho_1 > \rho_2$, this means that atom 1 has greater fuzziness (more vibration or more uncertainty) than atom 2. There can be many different ways to compute a fuzzy molecular surface. We choose to express a fuzzy molecular surface by molecular surfaces using a set of spheres each of that encloses each atom with the same probability. Even though the probability of fuzzy molecular surface is not same as the probability of atoms, this method would be useful to represent the different fuzziness on the molecular surface.



(a) Original Gaussian Distributions



(b) Rescaled Gaussian Distributions

Figure 2: Gaussian Distributions

We want to find ρ_1 and ρ_2 for which the probabilities of level sets are the same (figure 2 (a)). ρ_1 and ρ_2 can be found from the α-cuts of rescaled Gaussian distributions with a certain value p (figure 2 (b)) using the Lemma below.

Lemma: If we rescale the amplitude of each atom's distribution to 1 and find level sets with the same cut, the level sets of the original distributions with rescaled cut have the same probability.

Proof: Let $e^{-\frac{1}{2}(MX)^T U_{eq}'^{-1}(MX)} = p \Rightarrow -\frac{(MX)^T I(MX)}{2u_{eq}} = \ln p \Rightarrow |MX|^2 = 2u_{eq} \ln \frac{1}{p}$

Let the square of displacement $\rho(p)^2 = |MX|^2$ and the square of standard deviation $\sigma^2 = u_{eq}$. Thus we can rewrite the above equation as:

$$\rho(p)^2 = 2u_{eq} \ln \frac{1}{p} \Rightarrow \rho(p) = \sigma \sqrt{2 \ln \frac{1}{p}} \tag{5}$$

Let $t = \sqrt{2 \ln \frac{1}{p}}$, then for any two atoms i and j,

$$\begin{aligned} P(|X| \leq \rho_i(p)) &= \int_{|X| \leq t\sigma_i} G_i(X) dX = k_i \\ P(|X| \leq \rho_j(p)) &= \int_{|X| \leq t\sigma_j} G_j(X) dX = k_j \end{aligned}$$

If a point X follows a Gaussian distribution, the probability that X lies within a certain number of standard deviations from its mean is the same without respect to its mean and standard deviation.[13] Because $G_i(X)$ and $G_j(X)$ are Gaussian distributions, k_i is equal to k_j. This means the probability that the center of atom i lies within a sphere with a radius $\rho_i(p)$ is the same as the probability that the center of atom j lies within a sphere with a radius $\rho_j(p)$. As we have already mentioned, the sphere with a radius $\rho_i(p)$ is the level-set of atom i, and its confidence level is the probability that a center of the atom lies within the sphere. □

4. FUZZY MOLECULAR SURFACE

We define a *p-probability sphere* for each atom as the smallest sphere that contains the center of that atom with a probability p. To compute the fuzzy molecular surface, we first construct p-probability spheres for all the atoms. The fuzzy molecular surface is computed as a collection of P-probability surfaces where $p \leq P \leq 1$. Each P-probability surface is the crisp molecular surface constructed for p-probability spheres of that molecule.

Let $M = \{S_1, ..., S_n\}$, be a set of spheres, where S_i is expressed as a pair $< C_i, r_i >$ (C_i is the center point of S_i, and r_i is the radius of S_i). Let us define $d(X, Y)$ to be the Euclidean distance between X and Y. The power of a point X with respect to a sphere S_i is defined as $D(X, S_i) = d^2(X, C_i) - r_i^2$. Let us define the extended-radius sphere for atom i with atom radius r_i and probe-radius R to be $\Psi_i =< C_i, \gamma_i >$, where $\gamma_i = r_i + R$. The p-probability sphere for atom i that we defined at the beginning of this section can be expressed as $< C_i, \rho_i(p) >$. Let us further define the *extended-radius p-probability sphere* for atom i as the smallest sphere that contains the atom i with a probability p. It is defined as $\overline{\Psi}_i =< C_i, \gamma_i + \rho_i(p) >$. For the rest of this paper, we will use the overline symbol for the fuzzy atoms. We define the power of a point X with respect to a extended-radius p-probability sphere $\overline{\Psi}_i$ as:

$$\begin{aligned} D(X, \overline{\Psi}_i) &= d^2(X, C_i) - (\gamma_i + \rho_i(p))^2 \\ &= d^2(X, C_i) - \gamma_i^2 - (2\rho_i(p)\gamma_i + \rho_i(p)^2) \end{aligned}$$

By substitution of $\rho_i(p)$ from Equation (5),

$$D(X, \overline{\Psi}_i) = d^2(X, C_i) - \gamma_i^2 - (2\sigma_i\gamma_i\sqrt{2\ln\frac{1}{p}} + 2\sigma_i^2\ln\frac{1}{p})$$

Let $t = \sqrt{2\ln\frac{1}{p}}$ and $\rho_i(p) = \sigma_i\sqrt{2\ln\frac{1}{p}}$. Then,

$$D(X, \overline{\Psi}_i) = d^2(X, C_i) - \gamma_i^2 - (2\sigma_i t\gamma_i + \sigma_i^2 t^2) \tag{6}$$

Define a chordale $\overline{\Pi}_{ij}$ of two extended-radius p-probability spheres $\overline{\Psi}_i$ and $\overline{\Psi}_j$ as $\overline{\Pi}_{ij} = \{X | D(X, \overline{\Psi}_i) = D(X, \overline{\Psi}_j)\}$.

Now, $D(X, \overline{\Psi}_i) = D(X, \overline{\Psi}_j)$

$$\Rightarrow d^2(X, C_i) - \gamma_i^2 - (2\sigma_i t\gamma_i + \sigma_i^2 t^2) = d^2(X, C_j) - \gamma_j^2 - (2\sigma_j t\gamma_j + \sigma_j^2 t^2)$$

$$\Rightarrow 2(C_j - C_i)\cdot X + C_i\cdot C_i - C_j\cdot C_j + \gamma_j^2 - \gamma_i^2 - (\sigma_i^2 - \sigma_j^2)t^2 - 2(\sigma_i\gamma_i - \sigma_j\gamma_j)t = 0$$

Let $\Delta = (\sigma_i^2 - \sigma_j^2)t^2 + 2(\sigma_i\gamma_i - \sigma_j\gamma_j)t$. We know that a chordale of two spheres without atom fuzziness $\Pi_{ij} = 2(C_j - C_i)\cdot X + C_i\cdot C_i - C_j\cdot C_j + \gamma_j^2 - \gamma_i^2 = 0$. Thus, the equation of the chordale of two fuzzy spheres is:

$$\overline{\Pi}_{ij} = \Pi_{ij} - \Delta = 0 \Rightarrow (\sigma_i^2 - \sigma_j^2)t^2 + 2(\sigma_i\gamma_i - \sigma_j\gamma_j)t - \Pi_{ij} = 0 \tag{7}$$

To solve this equation, let us consider two cases when $\sigma_i^2 \neq \sigma_j^2$ and when $\sigma_i^2 = \sigma_j^2$.

Case A: $\sigma_i^2 \neq \sigma_j^2$,

$$t = \frac{\sigma_j\gamma_j - \sigma_i\gamma_i \pm \sqrt{(\sigma_i\gamma_i - \sigma_j\gamma_j)^2 + (\sigma_i^2 - \sigma_j^2)\Pi_{ij}}}{\sigma_i^2 - \sigma_j^2}$$

Therefore, the probability function for $\overline{\Pi}_{ij}$ at a point X is given by

$$F(X) = e^{-\frac{1}{2(\sigma_i^2 - \sigma_j^2)^2}(\sigma_j\gamma_j - \sigma_i\gamma_i \pm \sqrt{(\sigma_i\gamma_i - \sigma_j\gamma_j)^2 + (\sigma_i^2 - \sigma_j^2)\Pi_{ij}})^2}$$

If we let $A = \sigma_i^2 - \sigma_j^2$ and $B = \sigma_j\gamma_j - \sigma_i\gamma_i$, then

$$F(x) = e^{-\frac{1}{2A^2}(B \pm \sqrt{B^2 + A\Pi_{ij}})^2}$$

The probability distribution $F(X)$ has the greatest value when $B \pm \sqrt{B^2 + A\Pi_{ij}}$ is zero:

$$\begin{aligned} B &\pm \sqrt{B^2 + A\Pi_{ij}} = 0 \\ \Rightarrow\ & B^2 = B^2 + A\Pi_{ij} \\ \Rightarrow\ & A\Pi_{ij} = 0 \\ \Rightarrow\ & (\sigma_i^2 - \sigma_j^2)(2(C_j - C_i)\cdot X + C_i\cdot C_i - C_j\cdot C_j + \gamma_j^2 - \gamma_i^2) = 0 \end{aligned}$$

Which is the equation of $\Pi_{ij} = 0$. Thus, the probability distribution function $F(X)$ achieves its highest value on the chordale of the two crisp (non-fuzzy) atoms.

Case B: $\sigma_i^2 = \sigma_j^2$,

$$2(\sigma_i\gamma_i - \sigma_j\gamma_j)t - \Pi_{ij} = 0,\ t = \frac{\Pi_{ij}}{2(\sigma_i\gamma_i - \sigma_j\gamma_j)}.$$

Therefore, the probability function for $\overline{\Pi}_{ij}$ at a point X is given by

$$F(x) = e^{-\frac{1}{8B^2}\Pi_{ij}^2} \tag{8}$$

Combining the above cases we can say that the probability function for $\overline{\Pi}_{ij}$ is as follows:

$$F(x) = \begin{cases} e^{-\frac{1}{2A^2}(B \pm \sqrt{B^2 + A\Pi_{ij}})^2} & \text{for } \sigma_i^2 \neq \sigma_j^2 \\ e^{-\frac{1}{8B^2}\Pi_{ij}^2} & \text{for } \sigma_i^2 = \sigma_j^2 \end{cases} \tag{9}$$

Let us define the halfspace $\overline{H}_{ij}$ as $\overline{H}_{ij} = \{x | D(X, \overline{\Psi}_i) < D(X, \overline{\Psi}_j)\}$.

$$\begin{aligned} & D(X, \overline{\Psi}_i) < D(X, \overline{\Psi}_j) \\ \Rightarrow\ & D(X, \Psi_i) < D(X, \Psi_j) + \Delta \\ \Rightarrow\ & \Pi_{ij} < \Delta \ \text{ (From Equation (7))} \end{aligned}$$

Let us define the feasible cell, $\overline{F}_i$ for atom i as $\overline{F}_i = \bigcap_j \overline{H}_{ij}$.

The fuzzy molecular surface is computed from the feasible cells for fuzzy atoms described above using the algorithm presented by Varshney et al..[14]

The region of influence, $\rho_i(p)$, for atom i is defined by the fuzzy sphere $< C_i, r_i + \rho_i(p) + 2R + max_{j=1}^{n}(r_j + \rho_j(p)) >$ with the probability p, and the list of neighboring atoms, $\overline{N}_i$, for atom i is defined as $\overline{N}_i = \{j | d(C_i, C_j) < r_i + \rho_i(p) + 2R + r_j + \rho_j(p)\}$ or $\overline{N}_i = \{j | \overline{\Psi}_i \bigcap \overline{\Psi}_j \neq \phi\}$.

For each atom i, the feasible cell $\overline{F}_i = \bigcap_{j \in \overline{N}_i} \overline{H}_{ij}$ is computed by considering only its neighbors. After this, the closed components on $\overline{\Psi}_i$ are determined by finding the edges and faces of $\overline{F}_i$ that intersect $\overline{\Psi}_i$. Three kinds of surface patches, i.e. the convex spherical, concave spherical, and toroidal patches, can be generated from the closed components. The details follow Varshney et al.[14] and can be seen there.

5. IMPLEMENTATION AND RESULTS

5.1. Implementation overview

We have implemented a program for visualizing three-dimensional smooth molecular surface with and without atom vibrations and uncertainty. This program reads the information of each atom in a molecule such as geometric coordinates, atom type, and amino acid type, from a PDB file, then builds and visualizes a molecule. To express the fuzziness of the surface, the smooth molecular surface is visualized using multi-layered transparent surfaces, where the surface of each layer has a different confidence level and the transparency associated with the confidence level.

Figures 4 and 5 show the crisp molecular surfaces and corresponding fuzzy molecular surfaces. Each atom in a molecule has a different fuzziness, and the fuzzy molecular surface reflects the fuzziness of each atom. For example, in Figure 4 (b), the atoms at the upper left corner have greater fuzziness than the atoms at the upper right corner, which means that the atoms at the upper left corner move more or have more uncertainty. Therefore, the surface at the upper left corner is more cloudy than the surface at the upper right corner. To speed up the rendering time, less-detailed surfaces are used for the fuzzy surfaces so that smaller number of meshes can be used. Users can interactively change the maximum confidence level and the probe radius. The program is implemented on a Pentium4 1.5GHz PC with an NVIDIA Quadro2 Pro graphics card, using OpenGL Version 1.2.

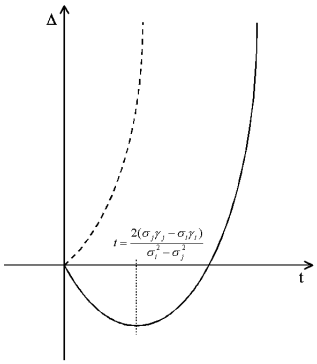


Figure 3: Displacement of a chordale

5.2. The change of a chordale with varying confidence level

$\Delta = (\sigma_i^2 - \sigma_j^2)t^2 + 2(\sigma_i\gamma_i - \sigma_j\gamma_j)t$ is the displacement of a chordale between two fuzzy atoms i and j from the chordale without atom fuzziness (See Section 4). If Δ is monotonous along with the change of t, the change of a chordale is also monotonous (the dotted line in figure 3). That means that the chordale moves in one direction. If the change of a chordale is not monotonous, the direction of the chordale movement varies (the solid line in figure 3). Δ is monotonous when $(\sigma_i^2 - \sigma_j^2)(\sigma_i\gamma_i - \sigma_j\gamma_j) \geq 0$ because Δ is a quadratic function and $t \geq 0$. This is shown as a dotted line in figure 3. If $(\sigma_i^2 - \sigma_j^2)(\sigma_i\gamma_i - \sigma_j\gamma_j) < 0$, the chordale moves in one direction when $0 < t \leq 2(\sigma_j\gamma_j - \sigma_i\gamma_i)/(\sigma_i^2 - \sigma_j^2)$, and moves in the opposite direction when $t > 2(\sigma_j\gamma_j - \sigma_i\gamma_i)/(\sigma_i^2 - \sigma_j^2)$. This is shown as a solid line in figure 3.

Therefore, chordales can be interpolated without computing them for all layers. We use this fact to speed up the generation of intermediate layers of the fuzzy molecular surface.

5.3. Results

Table 1: Running times with probe radius 1.4

Molecule	Surface type	Number of triangles	Construction time (sec)	Display time (sec)
Crambin (327 atoms)	Crisp Surface	27,359	0.490	0.050
	Fuzzy Surface	98,687	1.372	0.200
HIV Protease (760 atoms)	Crisp Surface	68,862	1.262	0.110
	Fuzzy Surface	239,262	4.015	0.451
Mycobacterium tuberculosis (1994 atoms)	Crisp Surface	153,448	3.935	0.230
	Fuzzy Surface	458,134	12.347	0.861

Table 1 shows the times taken for computing and visualizing the molecular surfaces. We used seven layers for the multi-layered surfaces including the crisp molecular surface. We use less-detailed surfaces for fuzzy surfaces, so the number of triangles for the fuzzy surface versus the number of triangles for the crisp surface is a factor between 3 and 3.6 instead of 7. To compute the intermediate layers, we interpolate the crisp surface and the highest layer whenever possible (as described in Section 5.2). The construction time for the fuzzy surface versus the construction time for the crisp surface is a factor between 2.8 and 3.2. The display time for the fuzzy surface is 3.7 to 4.1 times more than the display time for the crisp surface. This is more than what we would expect by just looking at the number of triangles. This is because the fuzzy layers use texture mapping and transparency which slow down their display.

6. CONCLUSIONS AND FUTURE WORK

In this paper, we have outlined an algorithm to compute and visualize a molecular surface from atoms with fuzzy positions using multi-layered surfaces. Using this algorithm, thermal motions and uncertainty of the

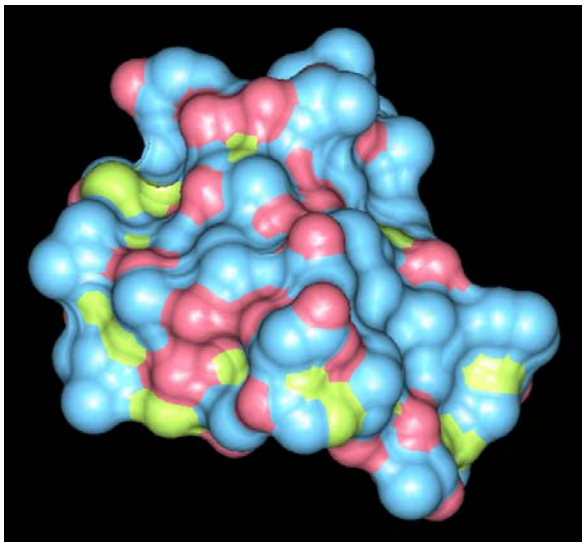
center positions of atoms can be represented in molecular surfaces. This provides a more informative molecular display that presumably lead to a better understanding of protein structure and function. We have assumed that the fuzziness of an atom to be isotropic, but a molecular surface with more general anisotropic displacement parameters could be more useful. We plan to consider this in future.

Acknowledgements

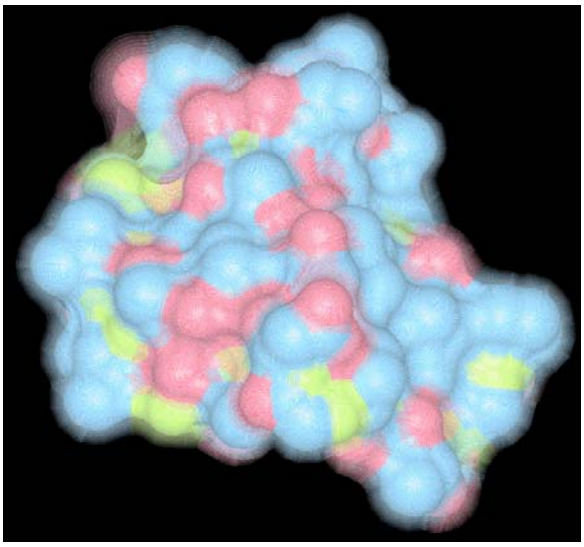
We would like to thank Azriel Rosenfeld for inspiring this work and John Moult for leading us in this direction. We would also like to acknowledge the reviewers for their very detailed and constructive comments which have led to a much better presentation of our results. This work has been supported in part by the NSF grants: IIS-00-81847, ACR-98-12572.

REFERENCES

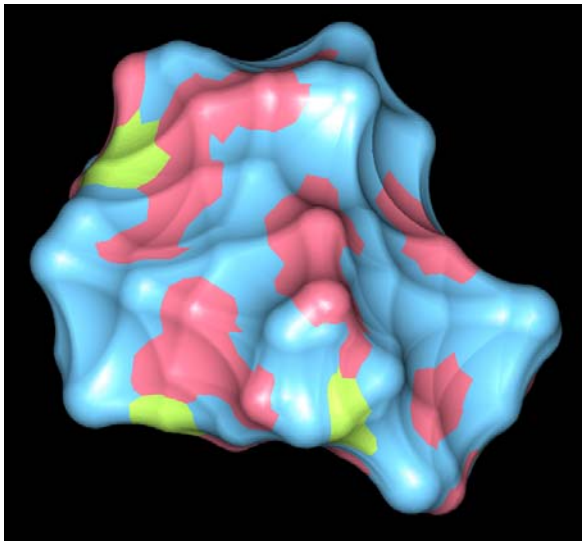
1. F. M. Richards, "Areas, volumes, packing and protein structures," in *Annual Review of Biophysics and Bioengineering*, **6**, pp. 151–176, 1977.
2. M. L. Connolly, "Analytical molecular surface calculation," in *Journal of Applied Crystallography*, **16**, pp. 548–558, 1983.
3. M. L. Connolly, "Solvent-accessible surfaces of proteins and nucleic acids," in *Science*, **221**, pp. 709–713, 1983.
4. M. L. Connolly, "Molecular surface triangulation," in *Journal of Applied Crystallography*, **18**, pp. 499–505, 1985.
5. A. Varshney, F. P. Brooks Jr., and W. V. Wright, "Linearly scalable computation of smooth molecular surfaces," *IEEE Computer Graphics and Applications* **14**, pp. 19–25, September 1994.
6. H. Edelsbrunner and E. P. Mucke, "Three-dimensional alpha shapes," *ACM Transactions on Graphics* **13**(1), pp. 43–72, 1994.
7. C. Bajaj, V. Pascucci, A. Shamir, R. Holt, and A. Netravali, "Multiresolution molecular shapes," in *TICAM Technical Report*, December 1999.
8. H. Huitema and R. van Liere, "Time critical computing and rendering of molecular surfaces using a zonal map," in *Virtual Environments*, pp. 115–124, 2000.
9. W. Cai, M. Zhang, and B. Maigret, "New approach for representation of molecular surface," *Journal of Computational Chemistry* **19**(16), pp. 1805–1815, 1998.
10. M. Sanner, A. Olson, and J. Spehner, "Reduced surface: an efficient way to compute molecular surfaces," *Biopolymers* **38**, pp. 305–320, 1996.
11. T. Klein, C. Huang, E. Pettersen, G. Couch, T. Ferrin, and R. Langridge, "A real-time malleable molecular surface," *Journal of Molecular Graphics* **8**(1), pp. 16–24 and 26–27, 1990.
12. A. Rosenfeld, "The fuzzy geometry of image subsets," in *Pattern Recognition Letters*, pp. 311–317, September 1984.
13. L. Ott, *An Introduction to Statistical Methods and Data Analysis*, PWS-Kent Publishing Company, 1988.
14. A. Varshney and F. P. Brooks Jr., "Fast analytical computation of Richards's smooth molecular surface," in *Proceedings of the IEEE Visualization*, pp. 300–307, (San Jose, CA), October 1993.



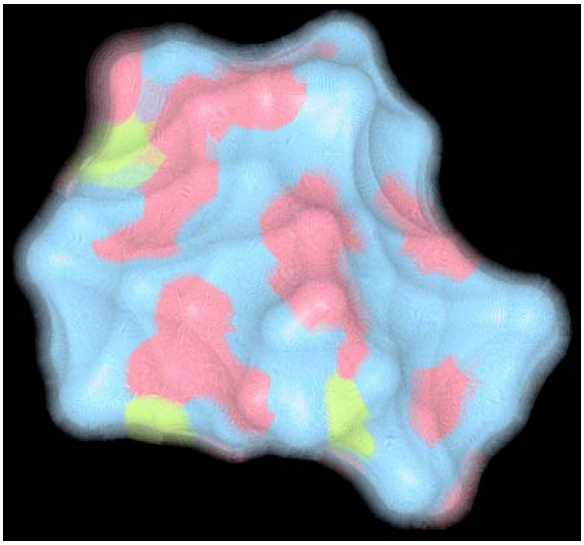
(a) Crisp molecular surface
(probe radius 1.4)



(b) Fuzzy molecular surface
(probe radius 1.4)

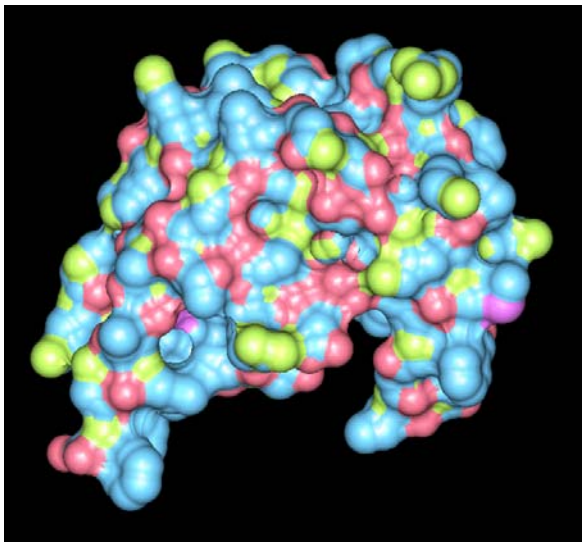


(c) Crisp molecular surface
(probe radius 5.0)

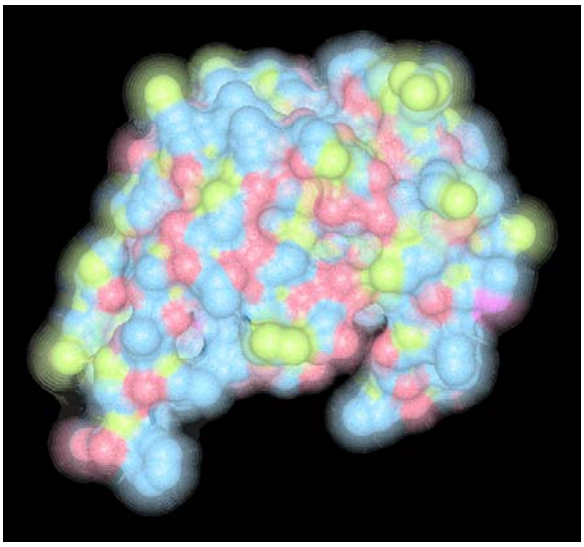


(d) Fuzzy molecular surface
(probe radius 5.0)

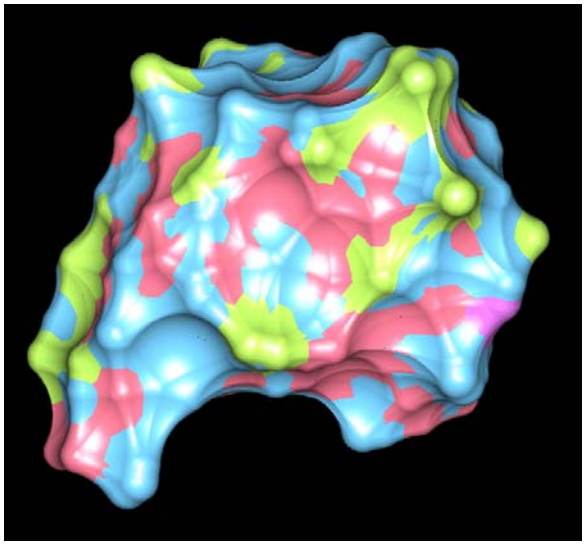
Figure 4: Molecular surfaces for Crambin



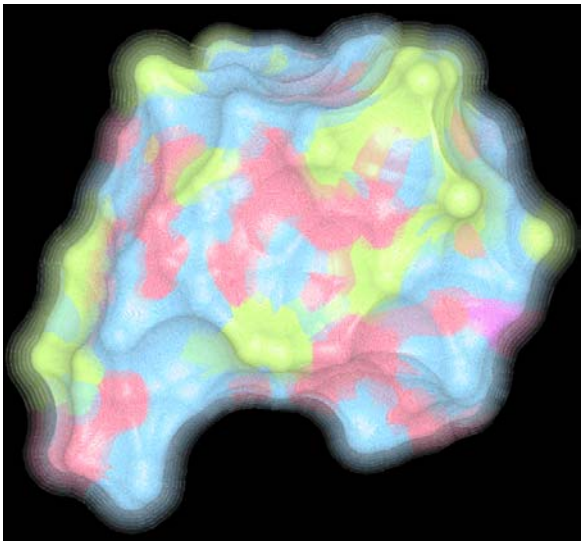
(a) Crisp molecular surface
(probe radius 1.4)



(b) Fuzzy molecular surface
(probe radius 1.4)



(c) Crisp molecular surface
(probe radius 5.0)



(d) Fuzzy molecular surface
(probe radius 5.0)

Figure 5: Molecular surfaces for HIV Protease