1

Geometric Processing of Reconstructed 3D Maps of Molecular Complexes

Chandrajit Bajaj
The University of Texas at Austin

Zeyun Yu
The University of Texas at Austin

1.1 Introduction

Today, hybrid experimental approaches for capturing molecular structures (henceforth, complexes), utilizing cryo-electron microscopy (cryo-EM), electron tomography (ET), X-ray crystallography (X-ray) or nuclear magnetic resonance spectroscopy (NMR), need to be ably complemented with faster and more accurate computational and geometric processing for ultrastructure elucidation at the best level of resolution that is possible [Fra96].

Electron Microscopy (EM) and in particular single particle reconstruction using cryo-EM, has rapidly advanced over recent years, such that several complexes can be resolved routinely at low resolution (10-20 $\mathring{A}$) and in some cases at sub-nanometer (intermediate) resolution (7-10 $\mathring{A}$) [BOF99]. These complexes provide not only insights into protein and nucleic acid folds, but perhaps even more importantly provide information about how the various structural components interact. There are increasing numbers of molecules where the tertiary or secondary structure of a complex can be fully determined using EM [ZBJ$^+$01]. Often the crystal structures of individual domains or components of these complexes are also known. An emerging trend in these fields is to fit the atomic resolution X-ray crystal structures into the cryo-EM map, to provide a quasi-atomic resolution model of the overall complex, possibly revealing details about molecular interactions within the assembly. In addition, with the increasing capability of determining multiple functional conformers of a complex, there is the promise of studying the dynamics of such interacting systems. The large physical size and complexity of such complexes combined with intermediate to low resolution models, presents challenges for structure to biological function determination.

0-8493-8597-0/01/$0.00+$1.50

This chapter reviews some of the crucial three dimensional geometric post-processing once a volumetric cryo-EM map (henceforth a 3D map) has been reconstructed, as essential steps towards an enhanced and automated computational ultrastructure determination pipeline. In particular the paper addresses 3D Map contrast enhancement, filtering, automated structural feature and subunit identification, and segmentation, as well as the development of quasi-atomic models from the reconstructed 3D Map via structure fitting.

1.2 Map Preprocessing

1.2.1 Contrast Enhancement

Many reconstructed 3D Maps, as well as captured 2D EM images, possess low contrast, or narrow intensity ranges i.e small differences between structural features and background densities, thereby making structure elucidation all the more difficult. Image contrast enhancement is a process used to "stretch" the intensity ranges, thereby improving the 2D image or 3D Map quality for better geometric postprocessing such as feature recognition, boundary segmentation, and visualization. The most commonly used methods in the past utilized global contrast manipulation based on histogram equalization [GW92, Pra91]. It is however well recognized today that using primarily global information is insufficient for proper contrast enhancement, as it often causes intensity saturation. Solutions to this problem include localized (or adaptive) histogram equalization [CLMS98, Sta00], which considers a local window for each individual image pixel and computes the new intensity value based on the local histogram defined within the local window. A more recently developed technique called the *retinex model* [JRW97b], in which the contribution of each pixel within its local window is weighted by computing the local average based on a Gaussian function. A later version, called the *multiscale retinex model* [JRW97a], gives better results but is computationally more intensive. Another technique for contrast enhancement is based on wavelet decomposition and reconstruction and has been largely used for medical image enhancement especially digital mammograms [LHW94, LSFH94].

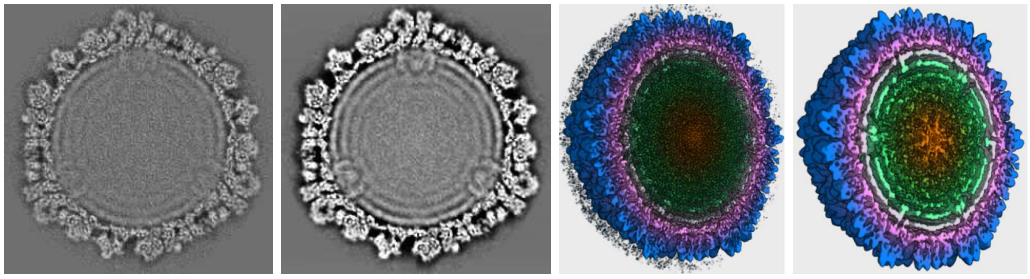
A fast and local method for 2D image or 3D Map contrast enhancement that we have obtained very good success with, is presented in [YB04a]. This is a localized version of classical contrast manipulations [GW92, Pra91]. The basic idea of this localized method is to design an adaptive one dimensional transfer function (mapping intensity ranges to intensity ranges) for each individual pixel (2D) or voxel (3D), based on the intensities in a suitable local neighborhood. There are three major steps, which we briefly describe for 2D images as its generalization to 3D Maps is straightforward. First, one computes local statistics (local average, minimum, and maximum) for each pixel using a fast propagation scheme [Der90, YV95]. The propagation rule from a pixel, say, $(m-1,\, n)$ to a neighboring pixel $(m,\, n)$ is defined as follows (similar propagation rules exist for other neighbors):

$$\text{lavg}_{m,n} = (1 - C) \times \text{lavg}_{m,n} + C \times \text{lavg}_{m-1,n} \tag{1.1}$$

where C is called the *conductivity factor*, ranging from 0 to 1. The matrix *lavg* stands for the local average map, initialized with the input image's intensity values. The above propagation rule is sequentially applied in row & column order [Der90, YV95]. In order to compute local min/max maps, some modifications are required for the above propagation scheme. To this end, a *conditional propagation scheme* is introduced in [YB04a]. Assume that *lmin* and *lmax* stand for the local min/max maps, respectively. The *conditional propagation scheme* from $(m-1,\, n)$ to $(m,\, n)$ is defined as follows:

$$\begin{cases} if(\text{lmin}_{m-1,n} < \text{lmin}_{m,n}) \\ \quad \text{lmin}_{m,n} = (1-C) \times \text{lmin}_{m,n} + C \times \text{lmin}_{m-1,n} \\ \\ if(\text{lmax}_{m-1,n} > \text{lmax}_{m,n}) \\ \quad \text{lmax}_{m,n} = (1-C) \times \text{lmax}_{m,n} + C \times \text{lmax}_{m-1,n} \end{cases} \tag{1.2}$$

Once these local statistics are calculated, the second step is to design the 1-dimensional adaptive transfer function, to achieve intensity range stretching on a per pixel basis. Similar to global contrast manipulations, various linear or nonlinear functions can be used here but all such functions should "extend" the narrow range of the local intensity histogram to a much broader range so as to achieve contrast enhancement. In the approach of [YB04a], the transfer function consists of two pieces: a convex curve (for stretching) in the dark-intensity range and a concave curve (for inverse stretching) in the bright-intensity range. The overall transfer function is C^1 continuous. Finally, in the last step, the intensity of each pixel is mapped to a new one using the calculated transfer function. This method inherits the advantages of the three afore-mentioned techniques, namely, global contrast manipulation, adaptive histogram equalization and the retinex model. However, unlike global contrast manipulation, this method is adaptive in the sense that the transfer functions are generally different from pixel to pixel. Also, unlike adaptive histogram equalization, this method considers a weighted contribution of each pixel within a local window. Furthermore, the size of the local window does not need to be pre-specified, due to the conditional propagation scheme used in this approach, which is also a significant difference between this method and the retinex model. Finally, the method of [YB04a] demonstrates a multi-scale property as different choosing different conductivity factors are chosen and used in the propagation scheme. Paper [YB04a], also gives an anisotropic version of the propagation scheme detailed above, and some results are shown in figure 1.1.



(a) (b) (c) (d)

FIGURE 1.1: Anisotropic filtering and contrast enhancement of the Rice Dwarf Virus (RDV). (a) original map (showing only one slice). (b) filtered and enhanced (same slice). (c) original map (volume-rendered). (d) filtered (volume-rendered).

1.2.2 Noise Reduction

Reconstructed 3D Maps are noisy due to both 2D image acquisition as well as computational errors in the 2D to 3D portion of the reconstruction pipeline [Fra96]. Applying 3D noise reduction techniques on the 3D maps as a pre-processing step, facilitates improved post-processing feature identification, segmentation and ultra structure determination. Traditional noise reduction filters applied to images include Gaussian filtering, median filtering, and frequency domain filtering [GW92]. Most of the recent research however, has been devoted to local anisotropic filters that operate with a directional bias, and vary in their ability to reduce noise without blurring the geometric structural features, especially edges and corners.

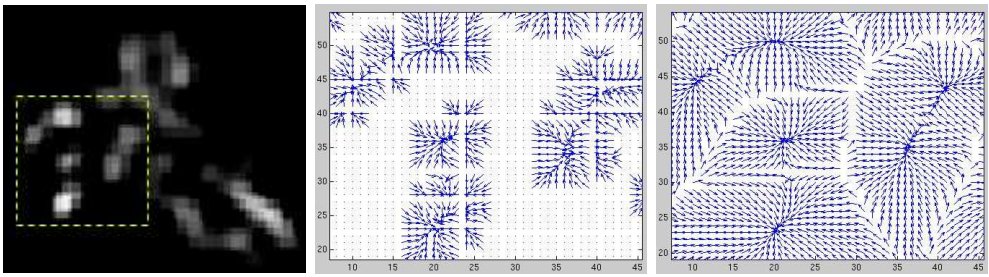
Bilateral filtering [Bar02, DD02, Ela02, TM98], or sometimes called weighted Gaussian filtering, uses an additional proximity weighting term to affect quasi-anisotropy. Partial differential equation (PDE) based filtering techniques, known popularly as anisotropic geometric diffusion [PM90a, Wei98a], differ primarily in the complexity of the local anisotropic modulation. Another popular anisotropic filtering approach is based on the use of the wavelet transformation [DJ94]. The basic idea is to identify and zero out wavelet coefficients of a signal that likely correspond to image noise while maintaining the sharpness of the edges in an image [XWHL94]. The development of nonlinear median-based filters in recent years has also produced promising results. One of these filters, the mean-median (MEM) filter [HLMR99, HK01], behaves differently from the traditional median filter, and has been shown to preserve fine details of an image while reducing noise. Among the aforementioned techniques, two noise reduction methods, namely wavelet filtering [SH97] as well as non-linear anisotropic diffusion [FH01], have also been applied to molecular tomographic imaging data.

An approach we have experimented successfully with on denoising reconstructed 3D maps, utilizes bilateral pre-filtering [JBW$^+$03], coupled to an evolution driven anisotropic geometric diffusion PDE (partial differential equation) [BWX03]. The PDE model is :

$$\partial_t \phi - \|\nabla\phi\| div\left(D^\sigma \frac{\nabla\phi}{\|\nabla\phi\|}\right) = 0 \tag{1.3}$$

The efficacy of our method is based on a careful selection of the anisotropic diffusion tensor D^σ based on estimates of the normal and principal curvature directions of a feature isosurface (level-set) in three dimensions [BWX03]. The diffusivities along the three independent directions of the feature boundary, are determined by the local second order variation of the intensity function, at each voxel. In order to estimate continuous first and second order partial derivatives, a tricubic B-spline basis is used to locally approximate the original intensity. A fast digital filtering technique based on repeated finite differencing, is employed to generate the necessary tri-cubic B-spline coefficients. The anisotropic diffusion PDE is discretized to a linear system by a finite element approach, and iteratively solved by the conjugate gradient method.

In Figure 1.1, we show an example of a reconstructed cryo-EM map and the results of filtering and contrast enhancement. In (a) and (b), only one slice of the 3D map is illustrated. In (c) and (d), a volume-rendering of the original map is compared to that of the filtered map.



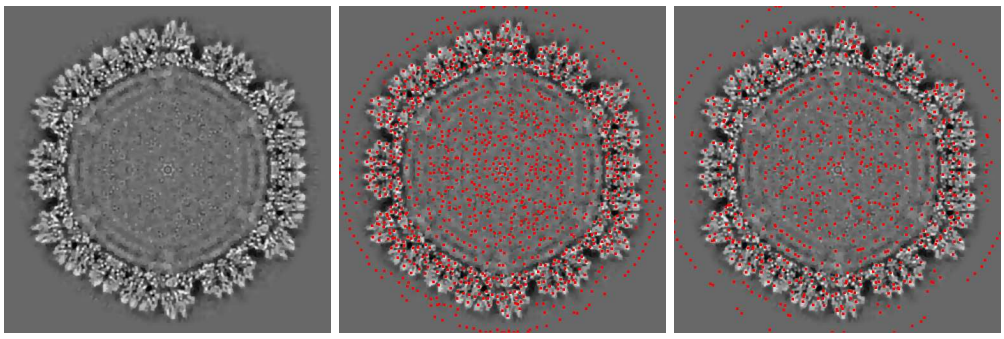
(a) (b) (c)

FIGURE 1.2: Illustration of critical point extraction using gradient vector diffusion. (a) one slice of herpesvirus capsid protein, vp5. (b) gradient vector field without diffusion corresponding to the boxed out area in (a). (c) gradient vector field after diffusion (10 iterations) improves the detection of critical points.

1.2.3 Gradient Vector Diffusion

In the earlier subsection we considered volumetric filtering in the special context of "critical" feature preservation. For a given volumetric map, the critical features are the essential values that help define the hierarchical structure of a complex. In general these critical features could be points, curves, or surfaces. The critical points of a scalar map can be classified as one of three types: local maxima, local minima, and saddle points of the given scalar function. However, in the context of structure identification, the maximal critical points are of great interest, due to the fact that, in a molecular density map, higher densities imply the existence of more atoms. These critical points can be easily computed from the local maxima of a given scalar map. Since noise is always present in the original maps, a pre-filtering process should be applied. As mentioned in the earlier subsection, a scalar map pre-filter can be either linear or nonlinear. A linear filter (e.g., Gaussian filtering) may destroy some weak features and hence eliminate some critical points. A nonlinear pre-filter [PM90b, Wei98b], however, tends to "deform" a sub-region, yielding many unwanted critical points.

A good alternative is a vector field filtering technique that is based on the diffusion of gradient vectors of the scalar 3D map, from which the afore-mentioned critical points are also easily extracted. In [XP98], the authors described a diffusion technique to smooth gradient vector fields. The gradient vectors are represented by Cartesian coordinates and a set of partial differential equations (PDEs) are separately applied to each component of the vectors. The equations are linear or isotropic, and therefore inherit the drawbacks of most linear filtering systems. A better way to diffuse a gradient vector field is based on the polar-coordinate representation of the vectors [YB02a, YB02b]. A drawback of this method is its computational complexity due to the efforts that have to be made to deal with the periodicity of orientation. An improved method is presented in [YB04b, BYA03], and we provide some details below.



(a) (b) (c)

FIGURE 1.3: Illustration of critical point extraction using gradient vector diffusion. (a) one slice (noise reduced) of rice dwarf virus (RDV). (b) after 10 iterations (1214 critical points were extracted). (c) after 30 iterations (781 critical points were extracted). The number of critical points can be further reduced by removing those whose density values are less than a certain threshold.

We detect the critical points using a set of anisotropic diffusion equations :

$$\begin{cases} \frac{du}{dt} = div(g(\alpha) \cdot \nabla u) \\ \frac{dv}{dt} = div(g(\alpha) \cdot \nabla v) \\ \frac{dw}{dt} = div(g(\alpha) \cdot \nabla w) \end{cases} \tag{1.4}$$

where (u, v, w) are initialized with the gradient vectors of the original maps. $g(\cdot)$ is a decreasing function and α is the angle between the central vector and its surrounding vectors. For instance, we can define $g(\alpha)$ as follows:

$$g(\vec{c}, \vec{s}) = \begin{cases} e^{\kappa \cdot (\frac{\vec{c} \cdot \vec{s}}{\|\vec{c}\| \|\vec{s}\|} - 1)} & if \ \ \vec{c} \neq 0 \ \ and \ \ \vec{s} \neq 0 \\ 0 & if \ \ \vec{c} = 0 \ \ or \ \ \vec{s} = 0 \end{cases} \tag{1.5}$$

where κ is a positive constant; $\vec{c}$ and $\vec{s}$ stand for the central vector and one of the surrounding vectors, respectively.

Once the gradient vector field is generated and diffused, we can define the *critical points* as those where none of the surrounding vectors is pointing away from those points. These critical points shall be frequently used in the following sections dealing with structural feature identification.

To better illustrate the application of the anisotropic gradient vector diffusion technique to accurately extract critical points from a given 3D map, we show cross-sectional two-dimensional (2D) slices (Figure 1.2). The images are from a slice of the herpesvirus capsid protein vp5 [ZDJ+00]. For better illustration of vector fields, we only consider a small area as boxed out in Figure 1.2(a). The vector field in Figure 1.2(b) is computed before the vector diffusion. Figure 1.2(c) demonstrates the power of the anisotropic vector diffusion, from which one can easily extract the critical points. Another example with greater detail,

is illustrated in Figure 1.3, where one can see that running the vector diffusion with different numbers of iterations can result in multiple levels of critical points.

1.3 Structural Feature Identification

1.3.1 Symmetry Detection

The symmetry of a shape or structure provides fundamental information for shape recognition, and representation. Given the reconstructed 3D map of a large biomolecular complex, one may ask: (1) Does this structure exhibit certain global and local symmetries ? (2) If it does, what type of symmetries are present (reflectional, rotational, translational, etc) ? (3) If the symmetry is rotational, what is the folding number and what is the location of the symmetry axis? Past relevant work devoted to answering the above questions in the literature include [SS97, MYY93, YLTL94, YC94, DG04, SICT99, LZ03], most of which, however, were applied to simpler inputs, such as a set of points, curves, or polygons.

In many cases, the 3D maps are of spherical viruses, whose protein capsid shells exhibit icosahedral symmetry. In these cases, the global symmetry detection can be simplified to computing the location of the 5-fold rotational symmetry axes, passing through the twelve vertices of the icosahedron, after which the 3-fold symmetry axis for the twenty icosahedron faces and the 2-fold symmetry axis for the thirty icosahedron edges can be easily derived. However local symmetries of the protein arrangement on virus capsid shells are more complicated, exhibiting varied k-fold symmetry and their detection requires a modified correlation based search algorithm explained below [YB05].

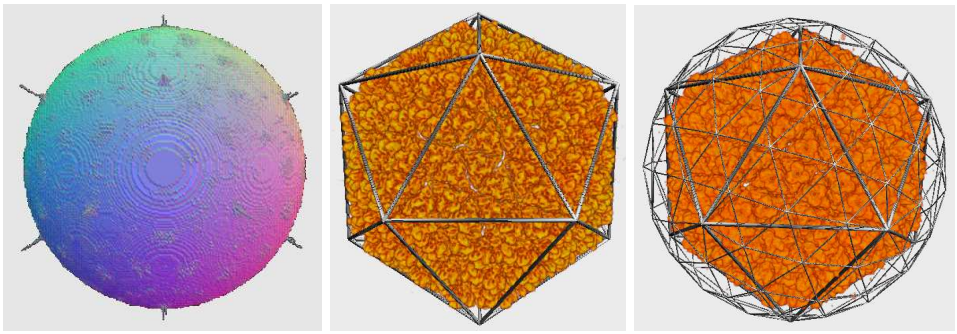
In almost all cases of single particle cryo-EM reconstruction, the origin of the 3D map is identical to the origin of its corresponding icosahedron, as global icosahedral symmetry is utilized in the reconstruction. Given an axis $l_{\theta,\varphi}$ passing through the origin, where θ and φ are defined in a classical way such that $\theta \in [-\pi, \pi]$ and $\varphi \in [-\pi/2, \pi/2]$, a 3D scalar map $f(\vec{r})$ is said to possess a 5-fold rotational symmetry about $l_{\theta,\varphi}$ if the following equation holds:

$$f(\vec{r}) = f(R_{(\theta,\varphi,2\pi/5)} \cdot \vec{r}), \text{ for } \forall \vec{r} \tag{1.6}$$

where the 3×3 matrix $R_{(\theta,\varphi,\alpha)}$ is defined as the coordinate transformation that rotates a point counterclockwise about an axis $l_{\theta,\varphi}$ by an angle of α. In particular, the matrix $R_{(\theta,\varphi,\alpha)}$ can be decomposed into five fundamental coordinate transforms.

In order to detect, for example a 5-fold symmetry axis, one can simply correlate the original map with its rotated map and search in the resulting correlation map for peaks [MYY93]. This method has a high computational complexity of $O(NM)$, where N is the number of voxels and M is the number of angular bins. In current applications of icosahedral virus reconstructions at medium resolution, N is roughly 700^3 and M is about 46,000 (a quasi-uniform sampling on the orientation sphere with a radius of 200-voxels). Although a number of techniques can be employed to speed up the search process by reducing the number of the angular bins (e.g., a principal axis method [SS97] or a coarse to fine hierarchical approach), it is still expensive as N is large. In prior recent work [YB04c, YB05], introduced a method for the fast detection of rotational symmetries, given the fold number. The idea there is to reduce N, the number of voxels to be tested, by restricting the correlation only to a subset of the critical points instead of the entire volume.

An example result of their method is shown in Figure 1.4(a) the scoring function of the outer capsid layer of the rice dwarf virus (RDV) 3D map [ZBJ+01], where one can clearly identify the "peaks" with high contrast. The corresponding 5-fold symmetry axes and the reconstructed icosahedra are shown in Figure 1.4(b). Experiments on this 3D Map



(a) (b) (c)

FIGURE 1.4: Detection of Symmetry axes and construction of global icosahedral symmetry as well as local n-fold symmetry. (a) scoring function. (b) global icosahedral symmetry. (c) local 6-fold symmetry.

data show that the correct symmetry axes could be calculated based only on 23 critical points, in contrast to the total number of 512^3 voxels in the original map (details are given in [YB04c, YB05]). The approach has been extended to automatically detect local symmetries, such as the 3- or 6-fold symmetry axes of the RDV map [YB05]. Figure 1.4(c) demonstrates the detection of the local symmetry axes of the outer capsid layer of RDV.

1.3.2 Boundary Segmentation

Segmentation is a way to electronically dissect significant biological components from a 3D map of a macromolecule, and thereby obtain a clearer view into the macromolecules architectural organization [Ell01]. For instance, it is often helpful to segment an asymmetric local subunit out of an icosahedral virus such that further structural interpretation can be conducted only on the asymmetric subunit instead of the entire map without loss of any structural information. Segmentation of 3D maps is usually carried out either manually [HYE96, KMM96, LLF97, MM99, HRS$^+$01] or semi-automatically [Vol02, FH02, BYA03]. Current efforts on the selection and decomposition of an icosahedral map into its local subunits also relies largely on manual work with extensive use of a graphical user interface [ZBJ$^+$01, JLB$^+$03]. This manual task can be tedious when the resolution is only marginally high enough to discern the boundaries between subunits.

Automated segmentation is still recognized as one of the challenge problems in image processing, although various techniques have been proposed for automated or semi-automated segmentation. Commonly used semi-automatic methods include segmentation based on edge detection, region growing and/or region merging, active curve/surface motion and model based segmentation (see for example [Set99, YB02b]). In particular, two techniques have been discussed in detail in the electron tomography community. One is called the *water-shed immersion method* [Vol02] and the other is based on normalized graph cut and eigenvector analysis [FH02].

Papers [YB04c, YB05] present steps towards an automatic approach for asymmetric subunit detection and segmentation of 3D maps of icosahedral viruses. The approach is an enhanced variant of the well-known fast marching method [MS98, Set96]. The basic idea of

the fast marching method is that a contour is initialized from a pre-chosen seed point, and the contour is allowed to grow until a certain stopping condition is reached. Every voxel is assigned with a value called *time*, which is initially zero for seed points and infinite for all other voxels. Repeatedly, the voxel on the marching contour with minimal *time* value is deleted from the contour and the *time* values of its neighbors are updated according to the following equation:

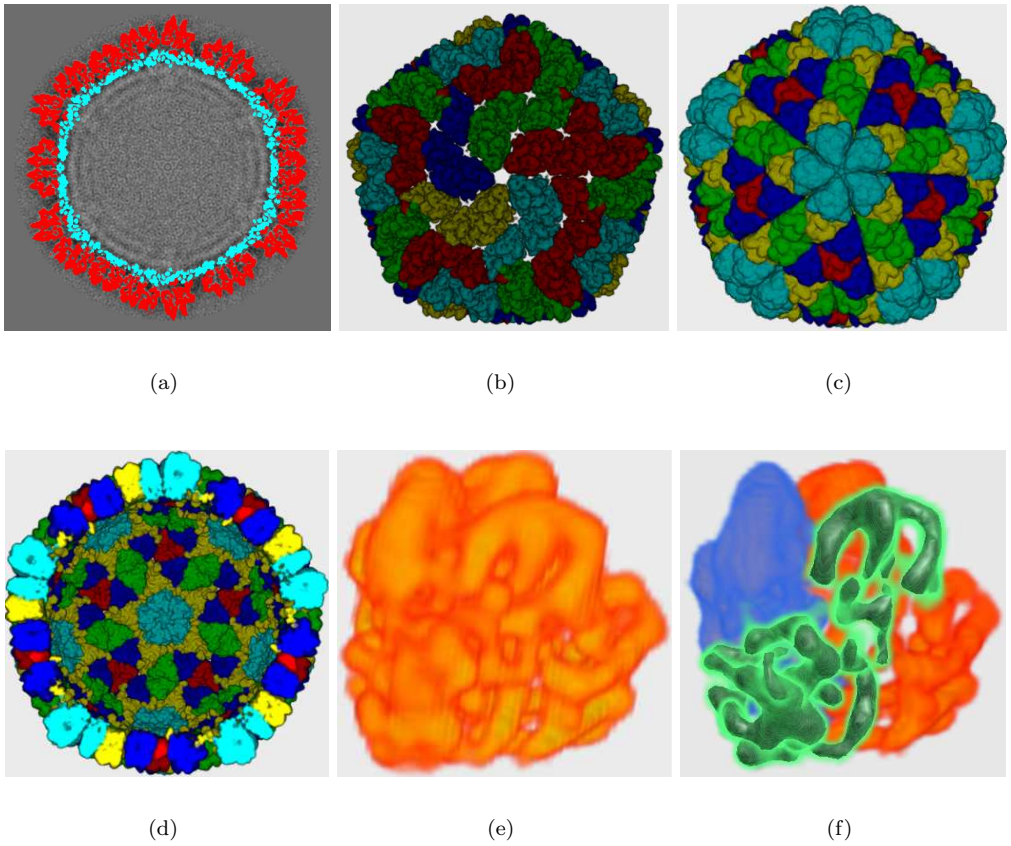
$$||\nabla T(\vec{r})|| \cdot F(\vec{r}) = 1 \tag{1.7}$$

where $F(\vec{r})$ is called the *speed function* that is usually determined by the gradients of the input maps (e.g., $F(\vec{r}) = e^{-\alpha\|\nabla I\|}$, where $\alpha > 0$ and I is the original map). The updated neighbors, if they are updated for the first time, are then inserted into the contour. The traditional fast marching method are designed for a single object boundary segmentation. In order to segment multiple targets, such as 60-component virus capsids or a 3-component molecular trimeric subunit, one has to choose a seed for each of the components. However, assigning only one seed to each component may cause appropriate boundary detection problems, as demonstrated in [YB05], and hence a *re-initialization* scheme becomes necessary.

The automatic approach of [YB04c, YB05] consists of three steps: (1) detection of the critical points; (2) classification of critical points; (3) a multi-seed fast marching method. The technique for (1) the detection of critical points has been briefly described in the earlier subsection on Gradient Vector Diffusion, of this chapter. All the critical points are regarded as seeds in the fast marching method. In general, the number of critical points in a map is much larger than the number of object components of interest. In other words, each component is assigned with a number of seeds instead of just one. Every seed initiates a contour and all contours start to grow simultaneously and independently. Two contours corresponding to the same component merge into a single contour, while two contours corresponding to different components stop on their common boundaries.

The initial classification of critical points as part of step (2) of the algorithm, is crucial in the segmentation of virus 3D maps. The critical points are classified utilizing local or global symmetry and based on their equivalence in terms of the asymmetric components that are to be segmented. Once all the seeds are classified, the above multi-seed variant of the fast marching method is used. First, each component initially possesses multiple seeds and hence multiple initial contours. Second, each marching contour is assigned a membership index based on the classification of seeds and the assignment to components. Once a voxel (volume element of the 3D Map) is conquered by a marching contour, it is assigned with the same index as the marching contour. Third, two marching contours with the same index merge into one when they meet, while two marching contours with different indices stop at their touching boundaries.

The segmentation approach or [YB04c] has been applied to the global asymmetric components dissection of icosahedral virus 3D maps. For viruses with more than 60 subunits that form a quasi-equivalent icosahedron, one additionally needs to incorporate the local symmetry axes of the viruses into the multi-seed classification and segmentation process[YB05]. Results from the above automatic segmentation technique applied to a reconstructed Cryo-EM 3D Map of the Rice Dwarf Virus (RDV)[ZBJ$^+$01] are shown in Figure 1.5. The RDV has double spherical protein shells (called capsids) with icosahedral symmetry. The first level segmentation is a separation of these two shells from the 3D map (see Figure 1.5 (a)). Next is a segmentation of the asymmetric subunits within each capsid. The sixty asymmetric subunits of the outer capsid viewed from the 5-fold symmetry axis is shown in Figure 1.5(b). Each subunit consists of four and one third trimeric sub-subunits [ZBJ$^+$01]. Figure 1.5 (c) and (d) illustrates the segmented trimers (260 in total), where (c) shows the view from outside while (d) shows the view from the inside. The segmentation shown in (c)



(a) (b) (c)

(d) (e) (f)

FIGURE 1.5: Visualization of the architecture of the Rice Dwarf Virus (RDV) 3D map (a) segmented outer and inner icosahedral capsid boundaries (b) segmented asymmetric subunits of the outer capsid (60 subunits in total). Each asymmetric subunit consists of four and one third trimers. (c) & (d) segmented trimeric subunits (260 in total), where (c) shows the view from the outside while (d) shows the view from inside. (e) each segmented trimeric subunit consists of three monomeric sub-subunits. (f) segmented monomeric subunit represents the 3D density map of a single P8 protein. The RDV 3D map data is courtesy Dr. Wah Chiu, NCMI,BCM, Houston

and (d) requires the local symmetry detection as shown in Figure 1.4(c) and the algorithm discussed in detail in [YB05]. Figure 1.5 (e) shows the segmented trimeric subunit consisting of three monomeric units, each of the same protein P8. Figure 1.5(f) shows the P8 protein monomeric unit segmented from the trimeric unit based on local 3-fold symmetry. It is worthwhile pointing out that in the visualization of the segmented trimeric subnits in Figure 1.5(b) only five colors are used to distinguish between sixty subunits, such that any five subunits surrounding the 5-fold symmetry axis would have different colors. In other directions, however, one may see two adjacent subunits having the same color although technically they have different component memberships. One can certainly find a more sophisticated coloring scheme to assure any two adjacent subunits always have different colors. Several more example segmentations for both reconstructed cryo-EM 3D Maps and

synthetic 3D maps generated from crystal structure data are given in [YB05].

1.3.3 Secondary Structure Identification

Although atomic structures are not detectable in reconstructed 3D cryo-EM maps, given their low feature resolution, it is sometimes feasible to locate secondary structures (alpha helices and beta sheets) from those maps [ZBJ+01, CBWZ02]. An approach for detecting alpha helices in 3D maps has been described in [WMSC01], where the alpha helix is modelled with a cylinder (length and thickness) and the cylinder is correlated with the segmented protein map. Since the best solution is achieved by exhaustively searching in translation space (3D) and orientation space (2D), this method is computationally expensive. In addition, this approach is designed only for alpha helix detection, not for the beta sheets. Another approach, designed for beta sheet detection, was recently proposed by [KM03, KXTM04]. This method uses a disk (planar) model for beta sheets. It inherits the disadvantage of slow computational speed due to the exhaustive search in both translation and orientation space, and furthermore cannot find curved beta sheets.

It is of course possible to combine the two methods above to detect both alpha helices and beta sheets, however to detect secondary structures efficiently one must avoiding the exhaustive search in both translation and orientation space. One possible approach is to consider scoring candidate helices/sheets only at the critical points of the 3D Map. This way, the search in translation space can be reduced to a significantly smaller number of locations. In addition, the search in orientation space at each critical point can be further reduced by utilizing the local structure tensor [FL03, Wei98a]. Given the 3D map $f(x, y, z)$, the gradient tensor is defined as:

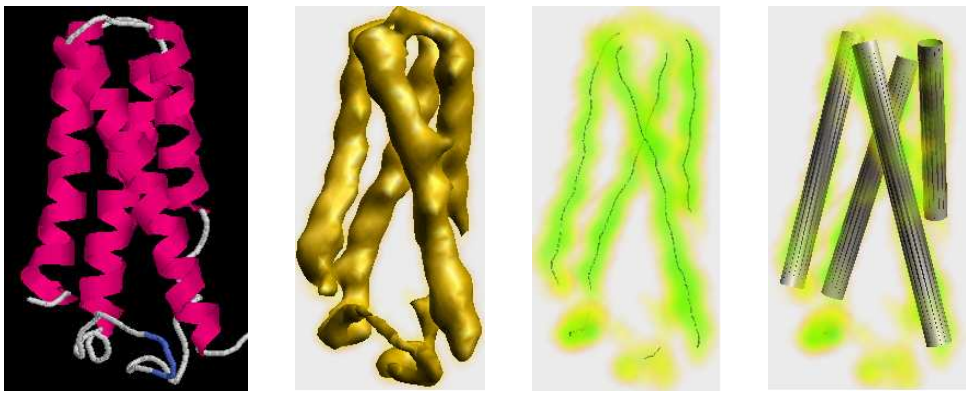
$$G = \begin{pmatrix} f_x^2 & f_x f_y & f_x f_z \\ f_x f_y & f_y^2 & f_y f_z \\ f_x f_z & f_y f_z & f_z^2 \end{pmatrix} \tag{1.8}$$

This matrix has only one non-zero eigenvalue: $f_x^2 + f_y^2 + f_z^2$. The corresponding eigenvector of this eigenvalue is exactly the gradient (f_x, f_y, f_z). Therefore, this matrix alone does not give more information than the gradient vector. To make the gradient tensor useful, a spatial average (over the image domain) should be conducted for each of the entries of the gradient tensors, yielding what is called the *local structure tensor*. The averaging is usually based on a Gaussian filter:

$$T = G_\alpha = \begin{pmatrix} f_x^2 * g_\alpha & f_x f_y * g_\alpha & f_x f_z * g_\alpha \\ f_x f_y * g_\alpha & f_y^2 * g_\alpha & f_y f_z * g_\alpha \\ f_x f_z * g_\alpha & f_y f_z * g_\alpha & f_z^2 * g_\alpha \end{pmatrix} \tag{1.9}$$

Here g_α is a Gaussian function with standard deviation α. The eigenvalues and eigenvectors of the structure tensor T indicate the overall distribution of the gradient vectors within the local window, similar to the well-known principal component analysis (PCA). Three typical structures can be characterized based on the eigenvalues [FL03]. Let the eigenvalues be $\lambda_1, \lambda_2, \lambda_3$ and $\lambda_1 \geq \lambda_2 \geq \lambda_3$. Then we have the following classifications:

1. blobs: $\lambda_1 \approx \lambda_2 \approx \lambda_3 > 0$.



(a) (b) (c) (d)

FIGURE 1.6: Illustration of secondary structural identification using local structure tensor at critical points of the 3D Map (a) The X-ray atomic structure representation of cytochrome c' (PDB-ID = 1bbh). (b) The volumetric representation of a Gaussian blurred 3D map generated from the X-ray structure (c) The detected skeletons of the 3D map. (d) Four helices are finally constructed from the skeletons, while the two on the bottom are discarded as being too small for being helices

2. lines: $\lambda_1 \approx \lambda_2 >> \lambda_3 \approx 0$.

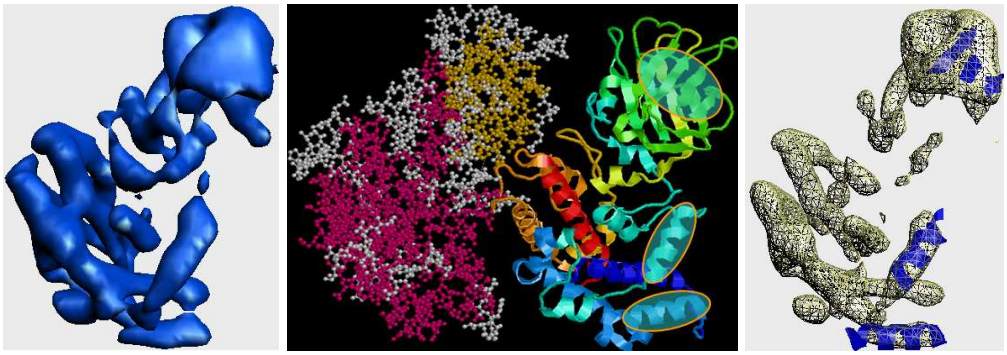
3. planes: $\lambda_1 >> \lambda_2 \approx \lambda_3 \approx 0$.

For each of the critical points of the 3D map, the structure tensor and its corresponding eigenvalues are calculated. Next, the above criterion based on the eigenvalues of the local structure tensor is computed at each of the critical points to distinguish between alpha helices (line features) and beta sheets (plane features). A critical point classified as an alpha helix, is to extended on both sides along the direction of the line structure determined by the local structure tensor, yielding a segment of the median axis of the 3D map. Similarly, for a critical point corresponding to a beta sheet feature, the plane feature is extended yielding a piece of median surface of the density map. Since a true alpha helix or beta sheet may consist of more than one critical point, it is necessary to merge a number of median segments and median surfaces, from which the final alpha helixes and/or beta sheets are constructed.

Figure 1.6 illustrates this approach on a Gaussian blurred map of the X-ray atomic structure of cytochrome c' (PDB-ID = 1bbh). Figure 1.6(a) shows the atomic structure, consisting of four alpha-helices, visualized as ribbons. The blurred map of this structure is visualized by contour rendering in Figure 1.6(b). Based on the the critical points of the 3D map and use of the structure tensor, the skeletons (median segments/planes) are computed and shown in Figure 1.6(c). From the skeletons, the four alpha helices are constructed as shown in Figure 1.6(d). Note that two segments of median axes on the bottom are discarded simply because their lengths are too small to be a true alpha helix.

1.4 Structure Fitting

A primary technique for structure interpretation and molecular model construction is to attempt to fit a known high-resolution structure (obtained by X-ray or NMR) into a reconstructed 3D density map. This technique is commonly known as *structure fitting* [Ros00b]. This technique bridges the resolution gap between low-resolution maps (e.g., lower than 10 Å) [BKF$^+$] and the atomic protein structures (e.g. lower than 3 Å). Figure 1.7 shows an example of structure fitting between the P8 monomeric protein, segmented from the RDV 3D map [ZBJ$^+$01], and its X-ray atomic structure [NMT$^+$03]. Figure 1.7(a) shows the segmented P8 monomeric protein (also see Figure 1.5(f)). The crystal structure of P8 monomer is shown in Figure 1.7(b), where one beta sheet (top) and two alpha helices (middle and bottom) are highlighted and used as a high-resolution fitting model. This high-resolution model is fit against the cryo-EM map of P8 monomer and its best position/orientation within the cryo-EM map is determined and show in Figure 1.7(c).



(a) (b) (c)

FIGURE 1.7: Example of structural fitting in the segmented P8 monomeric protein of RDV. (a) P8 monomeric protein iso-surface visualization (b) X-ray atomic structure of the P8 monomeric protein represented in ways of balls&sticks and cartoons. One beta sheet (top) and two alpha helices (middle and bottom) are highlighted and used as a fitting model. (c) By maximizing the correlation between the X-ray atomic model and the 3D map of the P8 monomer, one builds a pseudo-atomic model of the 3D map

There are several papers discussing various techniques on structure fitting. An excellent review of prior work on this topic is given in [WC01b]. One of the popular methods for volumetric matching is based on Fourier transforms [MM97, MR01]. The rigid-body fitting can be thought of as the minimization of the discrepancy between the cryo-EM maps and the atomic structure in Fourier space. The discrepancy is defined as follows:

$$R = \sum_{f} ||F_{em}(f)| - \lambda |F_{calc}(f,r,t)||^n, \qquad n = 1 \quad or \quad 2 \tag{1.10}$$

where F_{em} and F_{calc} are the Fourier transforms of the 3D map and the calculated atomic structure (that is, a Gaussian blurred 3D map of the atomic structure). Here r and t stand

for rotation and translation parameters, respectively, and both r and t have three degrees of freedom.

Instead of fitting the structures in Fourier space, we can also perform the fitting in the real space [VH99, KJ97, Cow98]. It is known that the minimization of the R factor seen above is equivalent to the maximization of the cross-correlation defined as below:

$$C_{=} \int \rho_{em}(\vec{x})\rho_{calc}(\vec{x}, r, t)d\vec{r} \tag{1.11}$$

where ρ_{em} and ρ_{calc} are the twin 3D maps of the cryo-EM and the Gaussian blurred atomic structure. The cross-correlation can be calculated by exhaustive searching with scaling or sampling of the translation (t) and rotation (r) parameters. While the Fast Fourier Transform (FFT) is easily used to speed up the cross-correlation scoring calculation over 3D translations [Cro72, KW02], it can also be used to compute the cross-correlation coefficients over rotational parameter (r) space, by first re-expressing the 3D map using trivariate spherical harmonics.

Another improvement on the conventional cross-correlation scoring method is to use a locally defined cross-correlation score [Ros00a]. In general, the global correlation method does not exclude the densities in the cryo-EM map that do not correspond to the atomic structure being considered. In addition, maximizing (1.11) often makes the solution "drift" to the highest density region in the cryo-EM map, which, however, does not mean the best-matched region. Hence the normalized and localized method [Ros00a] often gives more accurate fitting scores. One disadvantage of this method, however, is that the cross-correlation is conducted in real-space and a six-parameter searching space is considered in [Ros00a], resulting in a very slow performance. Recently, Roseman [Ros03] incorporated the fast Fourier transform (FFT) into the local correlation algorithm and applied it to the particle detection in two-dimensional electron micrographs. It was said that the local correlation algorithm together with FFT could be two orders of magnitude faster than the explicit real-space implementation [Ros03]. However, no results have been reported for 3D maps using this fast local correlation algorithm.

The conventional cross-correlation method can also be enhanced by a *contour-based fitting method* [CW01], in which the correlation coefficient is defined the same as (1.11) except that the Laplacian operator is applied to both maps before the calculation of the cross-correlation. Although this method is called *contour-based fitting*, it is not actually based on the detection of the contours. Due to the Laplacian operator that enhances the edges of both the cryo-EM map and the calculated atomic structure, this method was shown in [CW01] to give improved results (the resulting correlation map has higher contrast) than the classic cross-correlation method. However, as pointed out in [WC01b], the Laplacian filter may also amplify the noise, which as a result may weaken the performance of this method.

All the above methods for structure fitting are based on cross-correlation between the cryo-EM reconstructions and the calculated atomic structures. A different strategy is based on a data reduction technique. This method has been studied by Wriggers *et al* [WMSM98, WMM99, WC01a, WB01], based on a *vector quantization technique* [Gra83, IEE96]. The idea of vector quantization is to represent a 3D map with a certain number of vectors (or points in 3D space), from which a weighted graph is constructed. Instead of computing the cross-correlation between the cryo-EM 3D map and the calculated 3D atomic blurred map, one computes a new "difference" function between the two graphs corresponding to the cryo-EM map and the calculated atomic structure map. The "difference" function can be used to search for the best volumetric matching. Although this approach reduces the overall search time for its best match, and it is also possible to extend this to flexible fitting

[WC01b, WB01], it has two limitations. First, this method requires that the component of the cryo-EM map to be fitted should be isolated from the entire map. Second, the number of vectors must be carefully chosen. A large number of vectors exponentially increases the computational time while a small number of vectors may not be sufficient for perfect alignment and matching of the structural features of the map.

1.5 Conclusion

The field of structural biology, is increasingly dependent on computational processing for structural determination of complexes from 3D Maps. Each of the computational structure/ultra-structure elucidation methods that we highlighted above in separate subsections, remains an active area of future research and development, as there is still a ways to go. Nevertheless, we are optimistic that with progressively better techniques for image acquisition, coupled to efficient map reconstruction, and enhanced computational 3D map processing for structure elucidation, its only a matter of time when the resolution gap between X-ray structures and cryo-EM structures would be bridged.

References

References

[Bar02] D. Barash. A fundamental relationship between bilateral filtering, adaptive smoothing and the nonlinear diffusion equation. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 24(6):844–847, 2002.

[BKF+] D. M. Belnap, A. Kumar, J. T. Folk, T. J. Smith, and T. S. Baker. Low-resolution density maps from atomic models: How stepping.

[BOF99] T. S. Baker, N. H. Olson, and S. D. Fuller. Adding the third dimension to virus life cycles: three-dimensional reconstruction of icosahedral viruses from cryo-electron micrographs. *Microbiology and Molecular Biology Reviews*, 63(4):862–922, 1999.

[BWX03] C. Bajaj, Q. Wu, and G. Xu. Level-set based volumetric anisotropic diffusion for 3d image denoising. In *ICES Technical Report, University of Texas at Austin*, 2003.

[BYA03] C. Bajaj, Z. Yu, and M. Auer. Volumetric feature extraction and visualization of tomographic molecular imaging. *Journal of Structural Biology*, 144(1-2):132–143, 2003.

[CBWZ02] W. Chiu, M.L. Baker, J. Wen, and Z.H. Zhou. Deriving folds of macromolecular complexes through electron crymicroscopy and bioinformatics approaches. *Current Opin in Struct Biol*, 12:263–269, 2002.

[CLMS98] V. Caselles, J.L. Lisani, J.M. Morel, and G. Sapiro. Shape preserving local histogram modification. *IEEE Trans. Image Processing*, 8(2):220–230, 1998.

[Cow98] K. Cowtan. Modified phase translation functions and their application to molecular fragment location. *Acta Crystallography*, D54:750–756, 1998.

[Cro72] R. A. Crowther. The molecular replacement method. pages 173–178. Gordon & Breach, 1972.

[CW01] P. Chacon and W. Wriggers. Multi-resolution contour-based fitting of macromolecular structures. *Journal of Molecular Biology*, 317:375–384, 2001.

[DD02] F. Durand and J. Dorsey. Fast bilateral filtering for the display of high-

dynamic-range images. In *ACM Conference on Computer Graphics (SIGGRAPH)*, pages 257–266, 2002.

[Der90] R. Deriche. Fast algorithm for low-level vision. *IEEE Trans. on Pattern Recognition and Machine Intelligence*, 12(1):78–87, 1990.

[DG04] S. Derrode and F. Ghorbel. Shape analysis and symmetry detection in gray-level objects using the analytical fourier-mellin representation. *Signal Processing*, 84(1):25–39, 2004.

[DJ94] D.L. Donoho and I.M. Johnson. Ideal spatial adaptation via wavelet shrinkage. *Biometrika*, 81:425–455, 1994.

[Ela02] M. Elad. On the bilateral filter and ways to improve it. *IEEE Transactions On Image Processing*, 11(10):1141–1151, 2002.

[Ell01] R.J. Ellis. Macromolecular crowding: obvious but underappreciated. *Trends Biochem. Sci.*, 26(10):597–604, 2001.

[FH01] A. Frangakis and R. Hegerl. Noise reduction in electron tomographic reconstructions using nonlinear anisotropic diffusion. *J. Struct. Biol.*, 135, pages =, 2001.

[FH02] A. S. Frangakis and R. Hegerl. Segmentation of two- and three-dimensional data from electron microscopy using eigenvector analysis. *Journal of Structural Biology*, 138(1-2):105–113, 2002.

[FL03] J.-J. Fernandez and S. Li. An improves algorithm for anisotropic nonlinear diffusion for denoising cryo-tomograms. *J. Struct. Biol.*, 144(1-2):152–161, 2003.

[Fra96] J. Frank. *Three-dimensional Electron Microscope of Macromolecular Assemblies*. San Diego: Academic Press, 1996.

[Gra83] R.M. Gray. Vector quantization. *IEEE ASSP Mag.*, pages 4–29, 1983.

[GW92] R.C. Gonzalez and R.E. Woods. *Digital image processing*. Addison-Wesley, 1992.

[HK01] A. B. Hamza and H. Krim. Image denoising: A nonlinear robust statistical approach. *IEEE Transactions on Signal Processing*, 49(12):3045–3054, 2001.

[HLMR99] A. Ben Hamza, P. Luque, J. Martinez, and R. Roman. Removing noise and preserving details with relaxed median filters. *Journal of Mathematical Imaging and Vision*, 11(2):161–177, 1999.

[HRS+01] M.L. Harlow, D. Ress, A. Stoschek, R.M. Marshall, and U.J. McMahan. The architecture of active zone material at the frog's neuromuscular junction. *Nature*, 409:479 – 484, 2001.

[HYE96] D. Hessler, S. J. Young, and M. H. Ellisman. A flexible environment for the visualization of three-dimensional biological structures. *Journal of Structural Biology*, 116(1):113–119, 1996.

[IEE96] IEEE. Special issue on vector quantization. *IEEE Transactions on Image Processing*, 5(2), 1996.

[JBW+03] W. Jiang, M. Baker, Q. Wu, C. Bajaj, and W. Chiu. Applications of bilateral denoising filter in biological electron microscopy. *J. Struct. Biol.*, 144(1-2):114–122, 2003.

[JLB+03] W. Jiang, Z. Li, M. L. Baker, P. E. Prevelige, and W. Chiu. Coat protein fold and maturation transition of bacteriophage p22 seen at subnanometer resolution. *Nature Structural Biology*, 10(2):131–135, 2003.

[JRW97a] D.J. Jobson, Z. Rahman, and G.A. Woodell. A multiscale retinex for bridging the gap between color images and the human observation of scenes. *IEEE Trans. Image Processing*, 6(7):965–976, 1997.

[JRW97b] D.J. Jobson, Z. Rahman, and G.A. Woodell. Properties and performance of a

center/surround retinex. *IEEE Trans. Image Processing*, 6(3):451–462, 1997.

[KJ97] G. J. Kleywegt and T. A. Jones. Template convolution to enhance or detect structural features in macromolecular electron-density maps. *Acta Crystallography, D53*, pages 179–185, 1997.

[KM03] Y. Kong and J. Ma. A structural-informatics approach for mining b-sheets: locating sheets in intermediate-resolution density maps. *Journal of Molecular Biology*, 332:399–413, 2003.

[KMM96] J.R. Kremer, D.N. Mastronarde, and J.R. McIntosh. Computer visualization of three-dimensional image data using imod. *J Struct Biol*, 116:71–76, 1996.

[KW02] J. A. Kovacs and W. Wriggers. Fast rotational matching. *Acta Crystallography*, D58:1282–1286, 2002.

[KXTM04] Y. Kong, X.Zhang, T.S.Baker, and J. Ma. A structural-informatics approach for tracing b-sheets: building pseudo-ca traces for b-strands in intermediate-resolution density maps. *Journal of Molecular Biology*, 339:117–130, 2004.

[LHW94] J. Lu, D.M. Healy, and J.B. Weaver. Contrast enhancement of medical images using multiscale edge representation. *Optical Engineering*, 33(7):2151–2161, 1994.

[LLF97] Y. Li, A. Leith, and J. Frank. Tinkerbell-a tool for interactive segmentation of 3d data. *Journal of Structural Biology*, 120(3):266–275, 1997.

[LSFH94] A.F. Laine, S. Schuler, J. Fan, and W. Huda. Mammographic feature enhancement by multiscale analysis. *IEEE Trans. Medical Imaging*, 13(4):725–738, 1994.

[LZ03] G. Loy and A. Zelinsky. Fast radial symmetry for detecting points of interest. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 25(8):959–973, 2003.

[MM97] R. Mendelson and E. P. Morris. The structure of the acto-myosin subfragment 1 complex: results of searches using data from electron microscopy and x-ray crystallography. *Proc. Natl. Acad. Sci.*, 94:8533–8538, 1997.

[MM99] B.F. McEwen and M. Marko. Three-dimensional electron micros-copy and its application to mitosis research. *Methods Cell Biol*, 61:81–111, 1999.

[MR01] M. Mathieu and F. A. Rey. Atomic structure of the major capsid protein of rotavirus: implication for the architecture of the virion. *EMBO J.*, 20:1485–1497, 2001.

[MS98] R. Malladi and J. A. Sethian. A real-time algorithm for medical shape recovery. In *Proceedings of International Conference on Computer Vision*, pages 304–310, 1998.

[MYY93] T. Masuda, K. Yamamoto, and H. Yamada. Detection of partial symmetry using correlation with rotated-reflected images. *Pattern Recognition*, 26(8):1245–1253, 1993.

[NMT+03] A. Nakagawa, N. Miyazaki, J. Taka, H. Naitow, A. Ogawa, Z. Fujimoto, H. Mizuno, T. Higashi, Y. Watanabe, T. Omura, R.H. Cheng, and T. Tsukihara1. The atomic structure of rice dwarf virus reveals the self-assembly mechanism of component proteins. *Structure*, 11:1227–1238, 2003.

[PM90a] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 12(7):629–639, 1990.

[PM90b] P. Perona and J. Malik. Scale-space and edge detection using anisotropic diffusion. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 12(7):629–639, 1990.

[Pra91] W.K. Pratt. *Digital Image Processing (2nd Ed.)*. A Wiley-Interscience Pub-

lication, 1991.

[Ros00a] A. M. Roseman. Docking structures of domains into maps from cryo-electron microscopy using local correlation. *Acta Crystallographica, D56*, pages 1332–1340, 2000.

[Ros00b] M. G. Rossmann. Fitting atomic models into electron-microscopy maps. *Acta Crystallographica, D56*, pages 1341–1349, 2000.

[Ros03] A. Roseman. Particle finding in electron micrographs using a fast local correlation algorithm. *Ultramicroscopy*, 94:225–236, 2003.

[Set96] J. A. Sethian. A marching level set method for monotonically advancing fronts. *Proc. Natl. Acad. Sci.*, 93(4):1591–1595, 1996.

[Set99] J. A. Sethian. *Level Set Methods and Fast Marching Methods (2nd edition)*. Cambridge University Press, 1999.

[SH97] A. Stoschek and R. Hegerl. Denoising of electron tomographic reconstructions using multiscale transformations. *J. Struct Biol*, 120:257–265, 1997.

[SICT99] D. Shen, H. S. Ip, K. T. Cheung, and E. K. Teoh. Symmetry detection by generalized complex moments: a close form solution. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 21(5):466–476, 1999.

[SS97] C. Sun and J. Sherrah. 3d symmetry detection using the extended gaussian image. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 19(2):164–168, 1997.

[Sta00] J.A. Stark. Adaptive contrast enhancement using generalization of histogram equalization. *IEEE Trans. Image Processing*, 9(5):889–906, 2000.

[TM98] C. Tomasi and R. Manduchi. Bilateral filtering for gray and color images. In *1998 IEEE International Conference on Computer Vision*, pages 836–846, 1998.

[VH99] N. Volkmann and D. Hanein. Quantitative fitting of atomic models into observed densities derived by electron microscopy. *Journal of Structural Biology*, 125:176–184, 1999.

[Vol02] N. Volkmann. A novel three-dimensional variant of the watershed transform for segmentation of electron density maps. *Journal of Structural Biology*, 138(1-2):123–129, 2002.

[WB01] W. Wriggers and S. Birmanns. Using situs for flexible and rigid-body fitting of multiresolution single-molecule data. *J. Struct. Biol.*, 133:193–202, 2001.

[WC01a] W. Wriggers and P. Chacon. Using situs for the registration of protein structures with low-resolution bead models from x-ray solution scattering. *Journal of Applied Crystallography*, 34:773–776, 2001.

[WC01b] W. Wriggers and Pablo Chacon. Modeling tricks and fitting techniques for multiresolution structures. *Structure*, 9:779–788, 2001.

[Wei98a] J. Weickert. *Anisotropic Diffusion In Image Processing*. ECMI Series, Teubner, Stuttgart, ISBN 3-519-02606-6, 1998.

[Wei98b] J. Weickert. *Anisotropic Diffusion In Image Processing*. ECMI Series, Teubner, Stuttgart, ISBN 3-519-02606-6, 1998.

[WMM99] W. Wriggers, R. A. Milligan, and J. A. McCammon. Situs: a package for docking crystal structures into low-resolution maps from electron microscopy. *Journal of Structural Biology*, 125:185–195, 1999.

[WMSC01] J. Wen, M.L.Baker, S.J.Ludtke, and W. Chiu. Bridging the information gap: computational tools for intermediate resolution structure interpretation. *Journal of Molecular Biology*, 308:1033–1044, 2001.

[WMSM98] W. Wriggers, R. A. Milligan, K. Schulten, and J. A. McCammon. Self-organizing neural networks bridge the biomolecular resolution gap. *Journal of*

Molecular Biology, 284:1247–1254, 1998.

[XP98] C. Xu and J. L. Prince. Snakes, shapes, and gradient vector flow. *IEEE Trans. Image Processing*, 7(3):359–369, 1998.

[XWHL94] Y. Xu, J. B. Weaver, D. M. Healy, and J. Lu. Wavelet transform domain filters: A spatially selective noise filtration technique. *IEEE Trans. Image Processing*, 3(6):747–758, 1994.

[YB02a] Z. Yu and C. Bajaj. Anisotropic vector diffusion in image smoothing. In *Proceedings of International Conference on Image Processing*, pages 828–831, 2002.

[YB02b] Z. Yu and C. Bajaj. Image segmentation using gradient vector diffusion and region merging. In *Proceedings of International Conference on Pattern Recognition*, pages 941–944, 2002.

[YB04a] Z. Yu and C. Bajaj. A fast and adaptive algorithm for image contrast enhancement. In *Proceedings of International Conference on Image Processing*, pages 1001–1004, 2004.

[YB04b] Z. Yu and C. Bajaj. A segmentation-free approach for skeletonization of gray-scale images via anisotropic vector diffusion. In *Proceedings of 2004 IEEE International Conference on Computer Vision and Pattern Recognition*, pages 415–420, 2004.

[YB04c] Z. Yu and C. Bajaj. Visualization of icosahedral virus structures from reconstructed volumetric maps. *The University of Texas at Austin, Department of Computer Sciences. Technical Report TR-04-10*, 2004.

[YB05] Z. Yu and C. Bajaj. Automatic ultra-structure segmentation of reconstructed cryoem maps of icosahedral viruses. *IEEE Transactions on Image Processing: Special Issue on Molecular and Cellular Bioimaging (in review)*, 2005.

[YC94] K. Yuen and W. Chan. Two methods for detecting symmetries. *Pattern Recognition Letter*, 15:279–286, 1994.

[YLTL94] R. Yip, W. Lam, P. Tam, and D. Leung. A hough transform technique for the detection of rotational symmetry. *Pattern Recognition Letter*, 15:919–928, 1994.

[YV95] I.T. Young and L.J. Vliet. Recursive implementation of the gaussian filter. *Signal Processing*, 44:139–151, 1995.

[ZBJ+01] Z. H. Zhou, M. L. Baker, W. Jiang, M. Dougherty, J. Jakana, G. Dong, G. Lu, and W. Chiu. Electron cryomicroscopy and bioinformatics suggest protein fold models for rice dwarf virus. *Nature Structural Biology*, 8(10):868–873, 2001.

[ZDJ+00] Z. H. Zhou, M. Dougherty, J. Jakana, J. He, F. J. Rixon, and W. Chiu. Seeing the herpesvirus capsid at 8.5 . *Science*, 288:877–80, 2000.