Chapter 4. Clustering Core Atoms by Location

In this chapter, a process for sampling core atoms in space is developed, so that the analytic techniques in section 3C can be applied to local collections of core atoms. Localizing core atoms permits complex shapes to be assembled from simpler figures and permits multiple shapes to be differentiated from each other within a given image. It will be seen that the act of spatial sampling can, however, produce artifacts in the subsequent measurement of local medial dimensionality and orientation. This chapter explores these artifacts and develops a method to avoid them by clustering samples of core atoms that share a given medial neighborhood. The method is validated on 3D computer-generated shapes.

4A. Artifacts generated by Spatial Sampling

Core atoms can be analyzed in terms of their location if they are first sorted into bins on a regular 3D lattice. Since the location of a core atom is defined as the location of its center point, each bin represents a local sampling of medialness. The number of core atoms in each sample volume can be thought of as the medial *density* at that location.

How is an appropriate size for the sample volume to be selected? As will be seen, the local distribution of core atom centers can have a sizable cross section, and the density within that distribution is generally not uniform. Depending on the threshold used for face-to-faceness in selecting the core atoms, the diameter of the resulting *cloud* can be a considerable percentage of the object's physical cross-section. The diameter of the cloud is increased further because boundariness is itself probabilistic. Voxel intensity in a real image generally does not present the sharp boundaries of a physical object but rather a noisy transition in intensity and texture, which is detected by an aperture spanning multiple voxels. To preserve enough resolution to separate adjoining objects, the sample volume needs to be smaller than the typical cross-section of the core atom cloud. Therefore, the sample volume will generally not contain the entire thickness of the cloud.

Sampling a portion of the cloud results in distorted dimensionality, especially as one moves out from the center of the cloud. This concept is depicted in Figs. 4.1 and 4.2. The vector from the theoretical core (center of the sphere, axis of the cylinder) to the sample volume is

called the *displacement vector* $\vec{\mathbf{p}}$. As shown in Fig. 4.1, a zero-dimensional core at the center of a sphere will appear to be 1D (cylindrical) when sampled a bit off-center. Likewise, as shown in Fig. 4.2, the 1D core of a cylinder will appear to be 2D (slab-like). A sample of core atoms that is off-center is called a *coronal sample*.

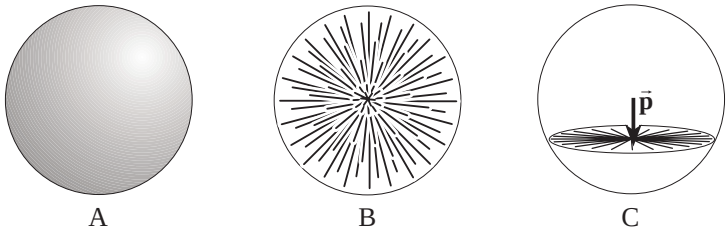

Fig. 4.1 **A.** Sphere. **B.** Core atom cloud. **C.** Sample displaced by $\vec{\mathbf{p}}$.

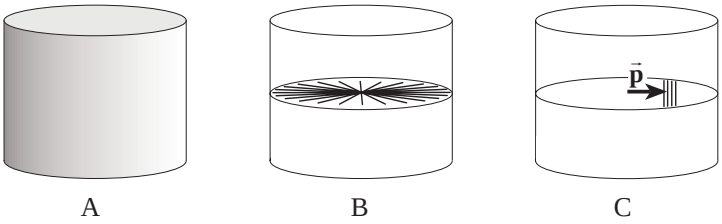

Fig. 4.2 **A.** Cylinder. **B.** Core atom cloud. **C.** Sample displaced by $\vec{\mathbf{p}}$.

The particular directions for $\vec{\mathbf{p}}$ in Fig. 4.1 and 4.2 are arbitrary. The center of the sphere in Fig. 4.1 is surrounded in all directions by cylindrical samples, each orthogonal a different displacement vector. The axis of the cylinder is surrounded in the same way by slab-like samples. These coronal samples comprise the thickness of the core atom cloud.

4B. Ellipsoidal Voting to Remove Sampling Artifacts

It is impossible to know the actual location of the center of the cloud from a single coronal sample, but multiple samples in the cloud can be combined to determine its center. Again, consider the shapes in Figs. 4.1 and 4.2. The population of core atoms in each coronal sample will be flattened in a plane orthogonal to $\vec{\mathbf{p}}$, and thus develop orthogonality to that direction. In section 3C eigenanalysis has been seen to determine the direction of maximum orthogonality

to a core atom population, namely the first eigenvector $\hat{\mathbf{a}}_1$ in Fig 3.8. The sample may use the eigenanalysis of its core atoms in a Hough-like fashion simultaneously to vote for itself and other samples whose corona it may inhabit. The voting takes place within ellipsoids around each sample. The axes of each ellipsoid are long in the direction(s) orthogonal to the core atoms in the sample, i.e., in the $\vec{\mathbf{p}}$ direction for that sample.

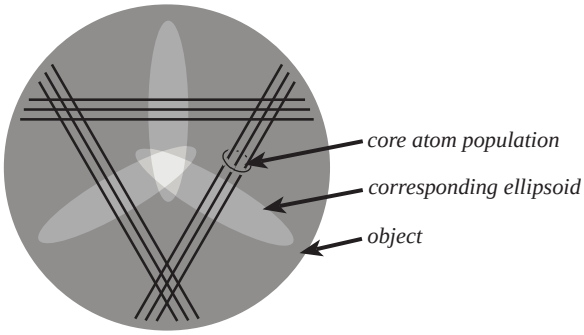

Fig. 4.3. Ellipsoids of three coronal core atom samples coalescing at the true center.

Fig. 4.3 demonstrates this concept. A circular cross-section through an object is shown with three coronal samples (each containing 3 core atoms) displaced from the center. An ellipsoid is associated with each sample, with the major axis of each ellipsoid oriented along the eigenvector most orthogonal to the corresponding core atoms. The three ellipsoids intersect at the center of the circle. Fig. 4.3 can be interpreted as the cross-section of a sphere with the populations of core atoms being cylindrical (seen in cross-section) and the ellipsoids intersecting at the center of the sphere (as in Fig. 4.1). Alternatively it can be interpreted as the cross-section of a cylinder with the populations of core atoms being slab-like and the ellipsoids intersecting along the axis of the cylinder (as in Fig. 4.2).

Various ways of constructing such ellipsoids are possible. The following heuristic was chosen for simplicity. The axes of each ellipsoid are oriented along the eigenvectors of the corresponding sample's **C** matrix. The lengths a_i $(i = 1,2,3)$ of the ellipsoid's three axes are related to the respective eigenvalues λ_i as follows:

$$a_1 = \gamma \bar{c}, \quad a_2 = \frac{\alpha_2}{\alpha_1} a_1, \quad a_3 = \frac{\alpha_3}{\alpha_1} a_1, \qquad \text{where } \alpha_i = 1 - \lambda_i, \quad \gamma = 1/2 \tag{4.1}$$

The scalar distance $\bar{c}$ is the mean diameter of the core atoms in the sample, and the dimensionless number γ relates $\bar{c}$ to the size of the ellipsoid, determining the general extent of overlap between ellipsoids.

The ellipsoids make it possible to cluster the core atoms in a given cloud, in effect, to coalesce the corona. Each sample (the *votee*) receives votes from all neighboring samples whose ellipsoids overlap it. The votes from a sample are assigned a strength v, where $v = n \cdot \exp\left(-d_e^{\,2}\right)$, n being the number of core atoms in the voting sample, and d_e being an *affine distance* from the center of the voter ellipsoid to the votee in coordinate system whose unit vectors are the axes of the ellipsoid $a_i\hat{\mathbf{a}}_i$ $(i = 1,2,3)$,

$$d_e = \sqrt{\sum_{i=1}^{3}\left(\frac{\hat{\mathbf{a}}_i \bullet \vec{\mathbf{d}}}{a_i}\right)^2} \quad , \tag{4.2}$$

where $\vec{\mathbf{d}}$ is the vector from voter to the votee.

Only samples containing core atoms are permitted to receive votes from their neighbors. This prevents spreading of votes to empty samples in the outer halves of ellipsoids if these regions are devoid of core atoms (see Fig. 4.3). Each vote is constructed to contain information about the constituent core atoms of the voter, including its **C** matrix. A weighted sum of the voters' **C** matrices is accumulated weighted by v for each voter and eigenanalysis performed yielding medial dimensionality and orientation of the core atom population in the votee's constituency. Thus are formed *clusters* of core atoms that no longer suffer from coronal distortion. Dimensionality, orientation, scale and center of mass after clustering tend to reflect the true core rather than any particular coronal sample. In the following section, this will be confirmed empirically on parametric test objects.

4C. Empirical Validation

Parametric Test Objects

The preceding concepts are demonstrated using three parametric test objects with simple geometries: the *sphere*, the *torus*, and the *spherical shell*. The torus is basically a cylinder of varied and known orientation, and the spherical shell is likewise a slab of varied and known orientation. The bin size was such that the major diameter of each shape represented the width

of between 10 and 15 bins. Fig. 4.4 shows the eigenvalues of all coronal samples containing greater than 1% of the entire core atom population, plotted on the lambda triangle. The arbitrary selection of 1% was to avoid samples with so few core atoms that their statistics would be meaningless. The sphere shows two groups of samples, one near the top (sphere) vertex of the triangle and another near the right (cylinder) vertex, consistent with the dimensional effects of the corona predicted in Fig. 4.1. The torus shows clustering near the right (cylinder) vertex, with some spreading towards the left (slab) consistent with the dimensional effects of the corona predicted in Fig. 4.2. The spherical shell shows tight clustering at the left (slab) vertex consistent with the observation that core atoms in a slab are collinear with $\vec{\mathbf{p}}$ and therefore do not develop dimensional distortion.

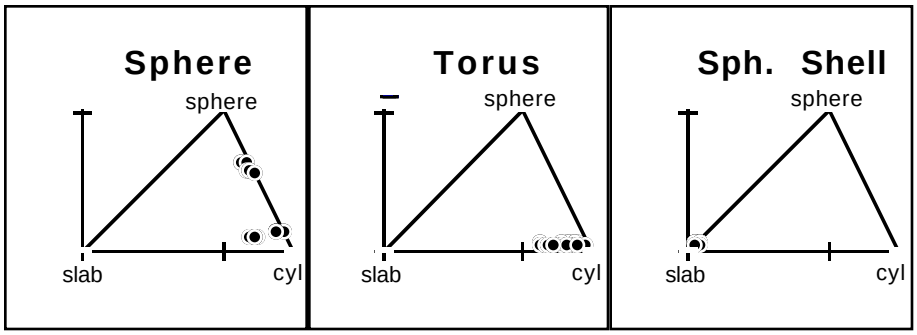

Fig. 4.4. Distribution of samples in lambda triangle for parametric test objects.

Spatial information about the samples for the same test objects is displayed in Fig. 4.5, which plots the number of core atoms in each sample as a function of displacement $\|\vec{\mathbf{p}}\|$ from the theoretical core. In each object the core atoms are concentrated as expected near the theoretical core, i.e., the center of the sphere, the axis of the cylinder, and the circle within the torus. Integer dimensionality (determined by the arbitrary partition of the lambda triangle as shown in Fig. 3.9) is labeled as follows: **o** = sphere, **x** = cylinder, **+** = slab. Dimensionality behaves as expected, clearly showing the predicted distortion with displacement from the theoretical core of the sphere and the cylinder. Also as expected, the slab (spherical shell) shows no such distortion.

The same core atom samples were analyzed in terms of their agreement with the theoretically expected values for medial orientation. One expects the displacement vector $\vec{\mathbf{p}}$ for a coronal sample to be one the eigenvectors at locations near the theoretical core because (1) by definition, the displacement vector will be orthogonal to the closest location on the core, and (2) the normal to the core is always one of its eigenvectors. In 3D, the medial manifold can

have at most 2 dimensions and thus will always have such a normal, unless it is itself a point in which case orthogonality is meaningless and the eigenvectors are arbitrary. A coronal sample cannot be a sphere.

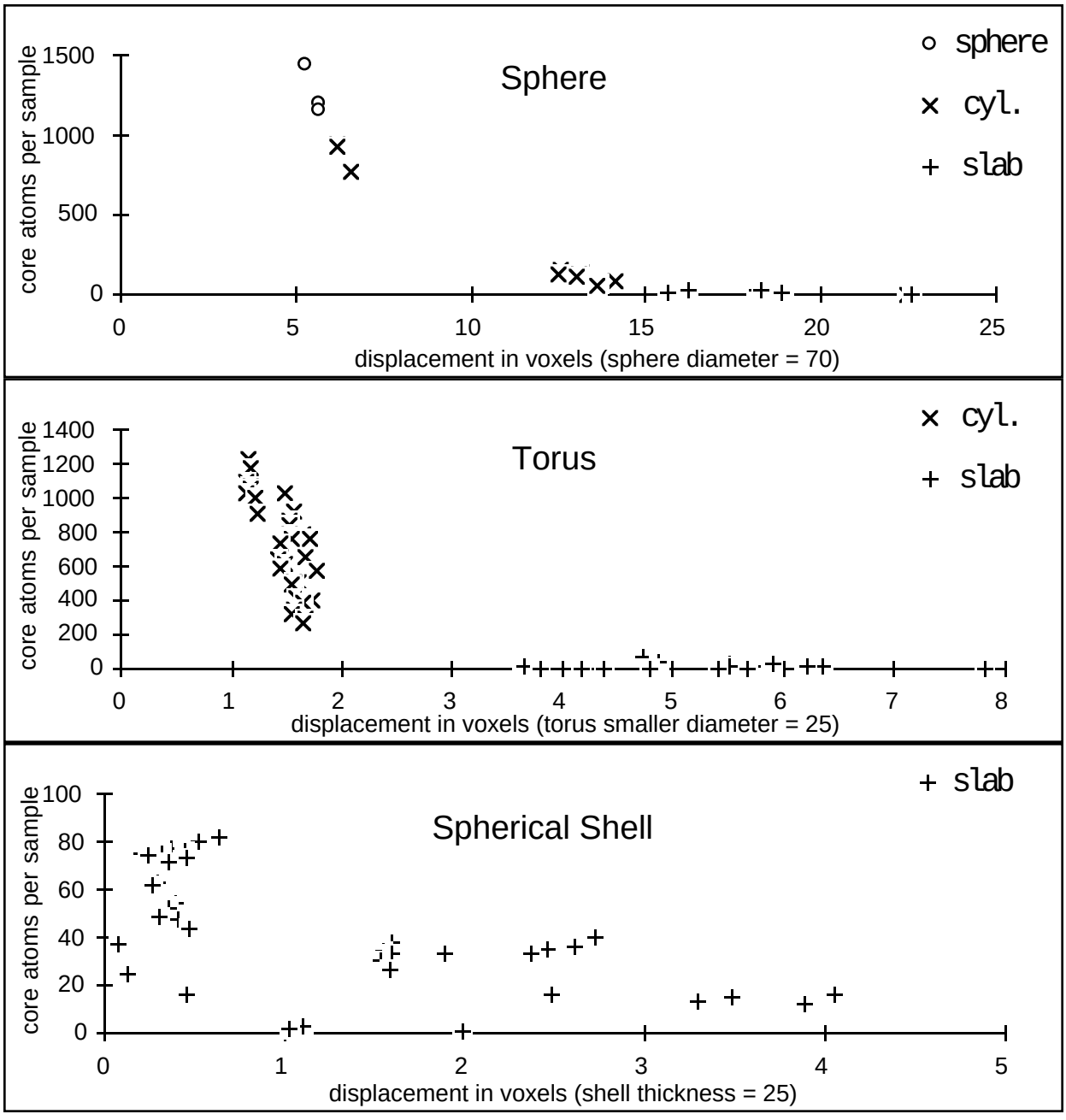

Fig. 4.5 Number of core atoms per sample vs. displacement from the theoretical core, showing dimensional distortion in the corona.

Since eigenanalysis of coronal samples can be expected to yield an eigenvector whose orientation nearly matches the displacement vector $\vec{\mathbf{p}}$, a measure of angular error ε_θ was used to select the closest eigenvector to $\hat{\mathbf{p}}$. That eigenvector was labeled the *displacement eigenvector* $\hat{\mathbf{a}}_p$, such that

$$\varepsilon_\theta = \arccos\left|\hat{\mathbf{a}}_i \bullet \hat{\mathbf{p}}\right|. \tag{4.3}$$

$$\hat{\mathbf{a}}_p = \arg\min_{\hat{\mathbf{a}}_i} \varepsilon_\theta \qquad i = 1,2,3 \tag{4.4}$$

For cylindrical samples, eigenanalysis of coronal samples can also be expected to yield an *axial eigenvector* $\hat{\mathbf{a}}_a$ whose orientation nearly matches the axis of the cylinder. Angular error ε_θ was between $\hat{\mathbf{a}}_a$ and the theoretical axis of the cylinder was determined in the case of the torus.

The mean angular error $\bar{\varepsilon}_\theta$ over the samples (weighted by density) for both $\hat{\mathbf{a}}_p$ and $\hat{\mathbf{a}}_a$ is reported in Table 4.1. The displacement eigenvectors $\hat{\mathbf{a}}_p$ are oriented with high accuracy for the sphere and the spherical shell, while the torus shows high accuracy in the alignment of the axial eigenvector $\hat{\mathbf{a}}_a$. The symmetry of the cylinder (torus) allows the $\hat{\mathbf{a}}_p$ to rotate freely at the core, which may explain the large $\bar{\varepsilon}_\theta$ for the $\hat{\mathbf{a}}_p$ of the torus.

MEAN ANGULAR ERROR $\bar{\varepsilon}_\theta$		
	$\hat{\mathbf{a}}_p$	$\hat{\mathbf{a}}_a$
sphere	6.1	-
torus	23.1	5.1
spherical shell	2.7	-

Table 4.1. Mean angular error $\bar{\varepsilon}_\theta$ (in degrees) for displacement eigenvector $\hat{\mathbf{a}}_p$, and axial eigenvector $\hat{\mathbf{a}}_a$ (cylinders only).

As one moves out along the displacement vector, the corresponding eigenvalue should drop towards zero as the sample develops orthogonality to $\vec{\mathbf{p}}$, except in the case of the slab, in which the core atom population can be expected to fall off rather abruptly without significant flattening. At the theoretical core, the eigenvalue in the direction of the displacement should approach 1/3 for a sphere, 1/2 for a cylinder, and 1 for a slab. These expectations were confirmed on the three test objects. Fig. 4.6 shows the eigenvalue associated with the displacement eigenvector $\hat{\mathbf{a}}_p$ as a function of displacement. In all cases (except a few aberrant

points for the torus) the eigenvalue behaved as expected, dropping to zero with increasing displacement for the sphere and torus, and with the value at the core approaching roughly $1/3$ for the sphere, $1/2$ for a cylinder, and 1 for a slab. As expected, the eigenvalue for the slab did not drop off with displacement. The aberrant points for the torus may have resulted from the symmetry of the cylinder, which, as noted above, allows $\hat{\mathbf{a}}_p$ to rotate more freely.

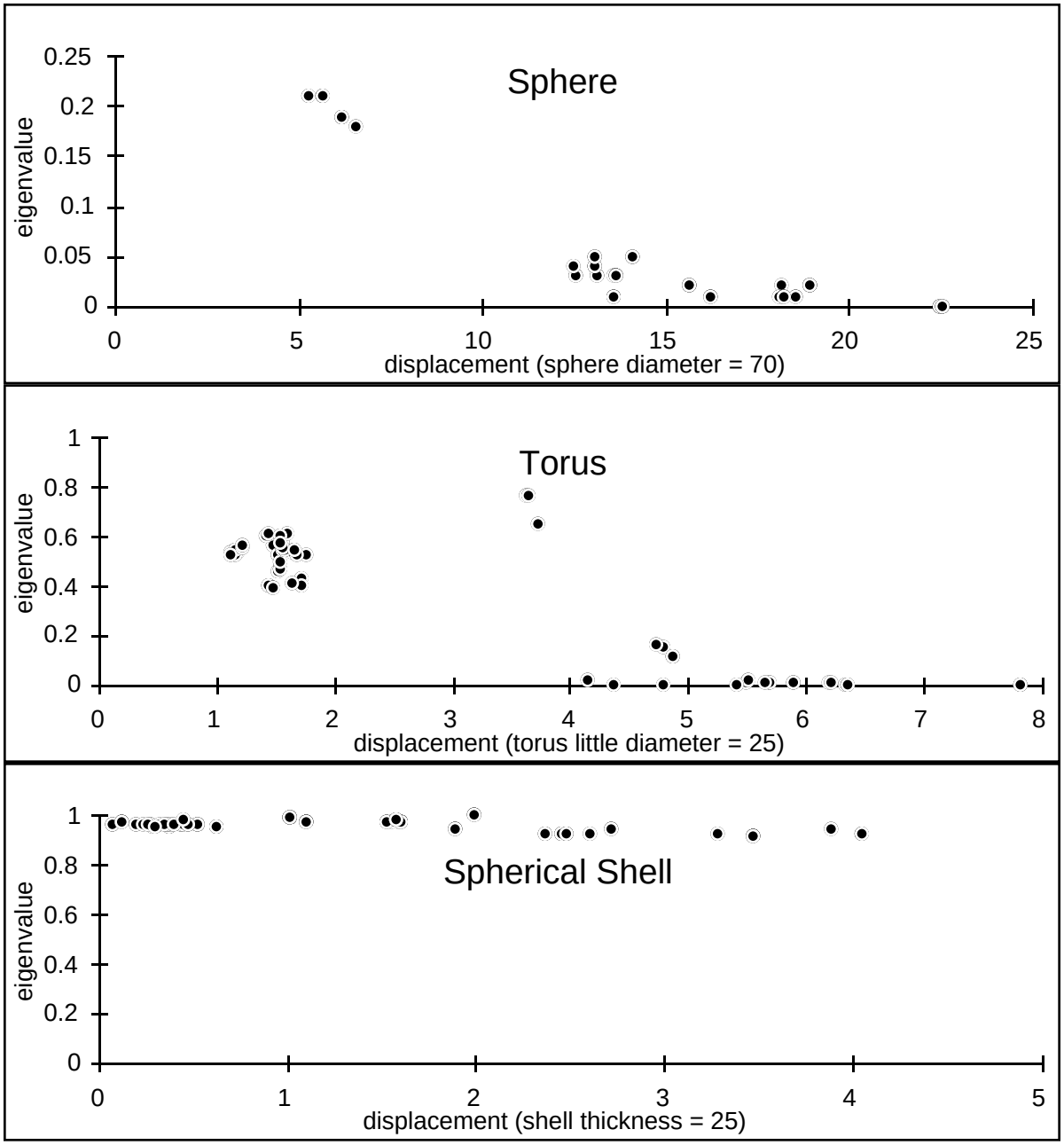

Fig. 4.6 Eigenvalue associated with eigenvector $\hat{\mathbf{a}}_p$ for all samples containing greater than 1% of the entire core atom population.

Ellipsoidal voting was applied to the core atom samples for each test object. Fig. 4.7 shows the samples for the sphere (dots). The sample that received the most votes was selected, and eigenanalysis of its cluster yielded eigenvalues closer to those expected for the sphere (labeled "+" 4.7) than and of the constituent samples. The center of mass of the cluster was also closer to the true center of the sphere than any of the samples. Further confirmation of the effectiveness of clustering by ellipsoidal voting is provided using 3D visualization techniques in the following section.

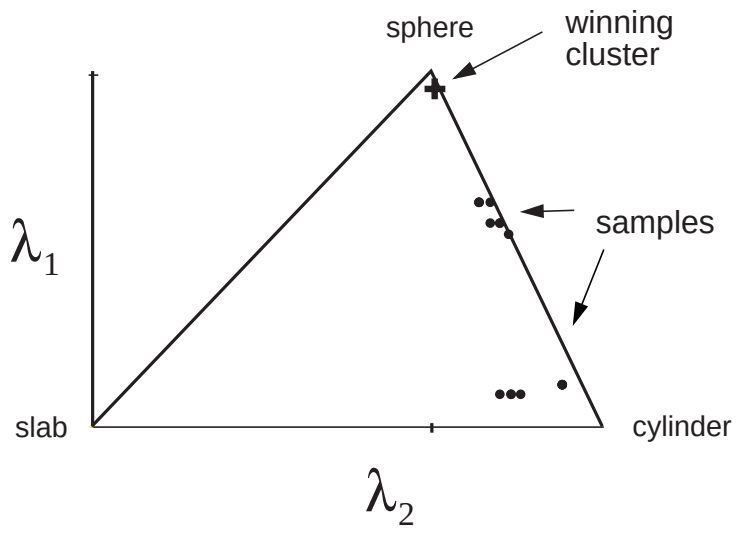

Fig. 4.7 Lambda triangle for the sphere showing dimensionality of samples (dots) and the cluster with the most votes (cross).

Visualizing the Corona

The spatial distribution of core atom samples is visualized in 3D in Fig. 4.8. Each sample volume that contains more than 1% of the total number of core atoms is shown as a thin-lined symbol. Here, the partition of the lambda triangle in Fig. 3.9 is used to decide between three possible symbols: a slab is represented as a single line, a cylinder as a cross, and a sphere as a small 3-axis symbol. The length of the thin lines in these test objects was purposely kept constant in order to maximize clarity. The orientation of the thin lines indicates the predominant direction(s) of core atoms in each sample, i.e., across the slab or orthogonal to the axis of the cylinder, keeping in mind that a perfect sphere has no predominant orientation and a perfect cylinder allows for arbitrary rotation around its axis.

As expected, the sphere shows cylindrical samples in its corona oriented towards its center. Further from the center a few slab-like samples reflect simply the paucity of core atoms in those samples. Near the center of the sphere, one true spherical sample (a small 3-axis symbol) can be discerned.

The thick-lined symbols show the results of ellipsoidal voting, i.e., they represent clusters of samples. To prevent a cluttered illustration, votes were tallied for all samples, a single winner was declared, and then all votes by that constituency for other candidates were removed. Using this Hough credit attribution strategy, the election was then repeated until all votes were gone. The winning clusters thus represent non-overlapping constituencies, making them easier to see in the illustration. The winning clusters are represented by thick lines in a manner similar to the initial samples, except the length of the axes now corresponds to the actual mean scale of the constituent core atoms. Thus the thick-lined 3-axis cross in the sphere indicates its diameter. For the sphere there is only one predominant winning cluster, with virtually every core atom in its constituency. In the corona of the sphere are clearly seen cylindrical samples oriented as expected to face the center of the sphere. The torus shows cylindrical initial samples properly oriented but dispersed throughout the corona. At the outer regions of the corona a few slab-like samples are visible. The clusters, by contrasts, are centered on the circular mid-line of the torus. The spherical shell shows only slab-like samples, which coalesce with ellipsoidal voting into slab-like clusters. The orientation of the initial samples and clustered samples are both across the local slab. Ellipsoidal voting is seen here to perform another function, that of connecting samples that share a core along the mid-plane of a slab or along the axis of a cylinder.

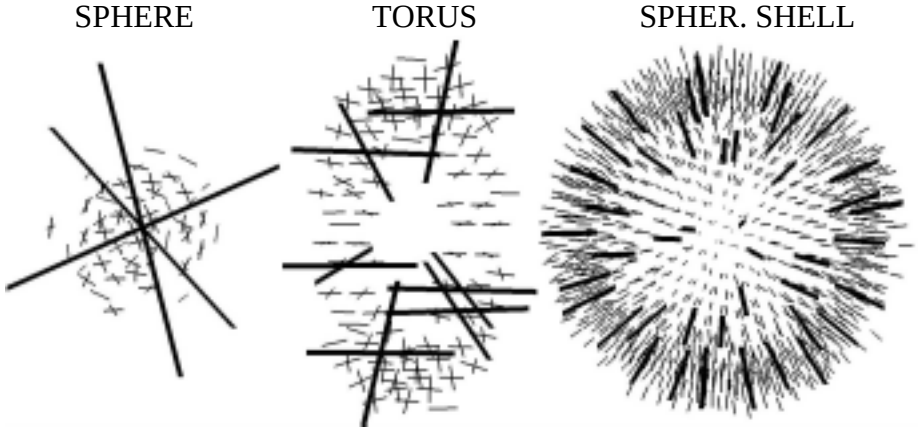

Fig. 4.8 Core atom samples (small symbols) and clustered samples (large symbols) for parametric objects (line = slab, cross = cylinder, 3-axis symbol = sphere).

If a color printer or display is available, the lambda triangle can be mapped to the standard Red-Green-Blue (RGB) color scheme (see Fig. 4.9) by assigning each vertex to a different color as follows:

slab = *red*
cylinder = *green*
sphere = *blue*

To ensure that the three corners of the lambda triangle are pure red, green and blue, the following formulae have been devised to calculate r, g, and b color values in the standard $[0,1]$ range used by the OpenGL graphics library,

$$\begin{aligned} r &= \tilde{n}^p(1-2\lambda_2)(1-3\lambda_1) \\ g &= \tilde{n}^p(2\lambda_2)(1-3\lambda_1) \\ b &= \tilde{n}^p(3\lambda_1) \end{aligned} \qquad (4.3)$$

where $\tilde{n}^p$ determines the *saturation* and the rest of each equation determines the *hue*. Since $0 \le \lambda_2 \le 1/2$ and $0 \le \lambda_1 \le 1/3$, locations in the lambda triangle along the λ_2 axis will vary from pure red to pure green. When $\lambda_1 = 1/3$ the hue will be pure blue. The color saturation from bright to dark depends on $\tilde{n}$, the number of core atoms in a sample (or cluster) normalized to the maximum number in that sample. The exponent p permits contrast to be adjusted between $p = 0$ (no differentiation by core atom number) to $p >> 0$ (only large numbers of core atoms visible). Using this scheme, color is mapped onto each sample and cluster in Fig. 4.10.

Measuring Scale

Besides dimensionality, orientation, and location, it is also possible to use local clusters of core atoms to measure medial scale. Each core atom contains a measurement of distance between its two boundary points. The mean length of a local population of core atoms is therefore a statistical measure of an object's diameter.

This is demonstrated on the same three test objects as before. A measurement of diameter was made for each object by averaging the diameter from all samples whose integer dimensionality matched that of the theoretical core. Thus only spherical samples were considered for the sphere, etc. This eliminated coronal core atoms, which do not pass through

the center of the object and whose length therefore tends to be shorter. These results are shown in the Table 4.2. As can be seen, an accurate determination of diameter was made for each object.

MEASUREMENT OF DIAMETER				
object	dim.	actual	measured	
			mean	s.d.
sphere	0	70	69.20	0.08
torus	1	25	23.90	0.07
sph. shell	2	25	22.30	1.50

Table 4.2 Actual diameter of test objects, and diameter determined from populations of core atoms (mean and s.d.)

The method of clustering samples of core atoms developed in this chapter has been shown to be capable of delivering local measurements of medial dimensionality, orientation, location, and scale. In the next chapter, these local properties will be used to construct more complicated shapes, and the question will be addressed of what happens at the end of a core.

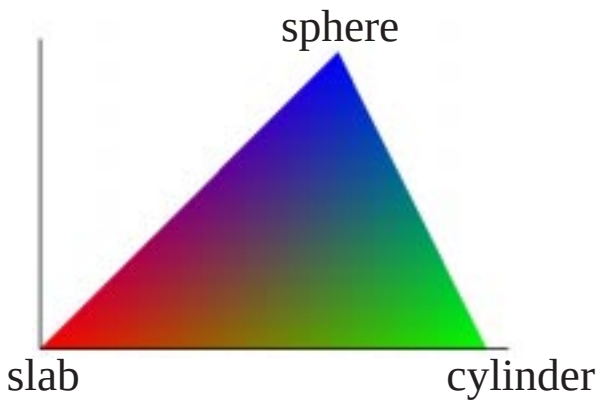

Fig. 4.9 Color mapped onto the lambda triangle.

Fig. 4.10 Color version of Fig. 4.8 with color representing dimensionality as mapped onto the labmda triangle in Fig. 4.9.