Chapter 5

The Spherical Harmonic Descriptor

5.1 Spherical harmonics

There are many possibilities for choosing a set of basis functions which are defined on the sphere. The spherical harmonic functions are a popular choice because they are relatively simple and have a number of nice mathematical properties. They are introduced e.g. by Greiner [18].

5.1.1 Definition

Y_l^m denotes the spherical harmonic function of degree l and order m. The following definitions agree with [31]. The variable w is a scalar and will correspond to u_2 below.

Legendre polynomials

$$P_l(w) = \frac{1}{2^l l!} \frac{d^l}{dw^l} (w^2 - 1)^l \tag{5.1}$$

Associated Legendre polynomials

$$\begin{aligned} P_l^m(w) &= (-1)^m (1 - w^2)^{\frac{m}{2}} \frac{d^m}{dw^m} P_l(w) \\ &= \frac{(-1)^m}{2^l l!} (1 - w^2)^{\frac{m}{2}} \frac{d^{m+l}}{dw^{m+l}} (w^2 - 1)^l \end{aligned} \tag{5.2}$$

$$Y_l^m(\theta,\phi) = \sqrt{\frac{2l+1}{4\pi}\,\frac{(l-m)!}{(l+m)!}}\;\; P_l^m(\cos\theta)\;\; e^{im\phi} \tag{5.3}$$

$$Y_l^{-m}(\theta,\phi) = (-1)^m Y_l^{m\,*}(\theta,\phi) \tag{5.4}$$

A list of the spherical harmonics up to degree 3 (in table 5.1) exemplifies these definitions. Both polar coordinates (θ,ϕ) and Cartesian coordinates (u_0,u_1,u_2) are used. The Cartesian notation reveals that spherical harmonics are just polynomials, in spite of the $\frac{m}{2}$ exponent, which means a square root if m is odd.

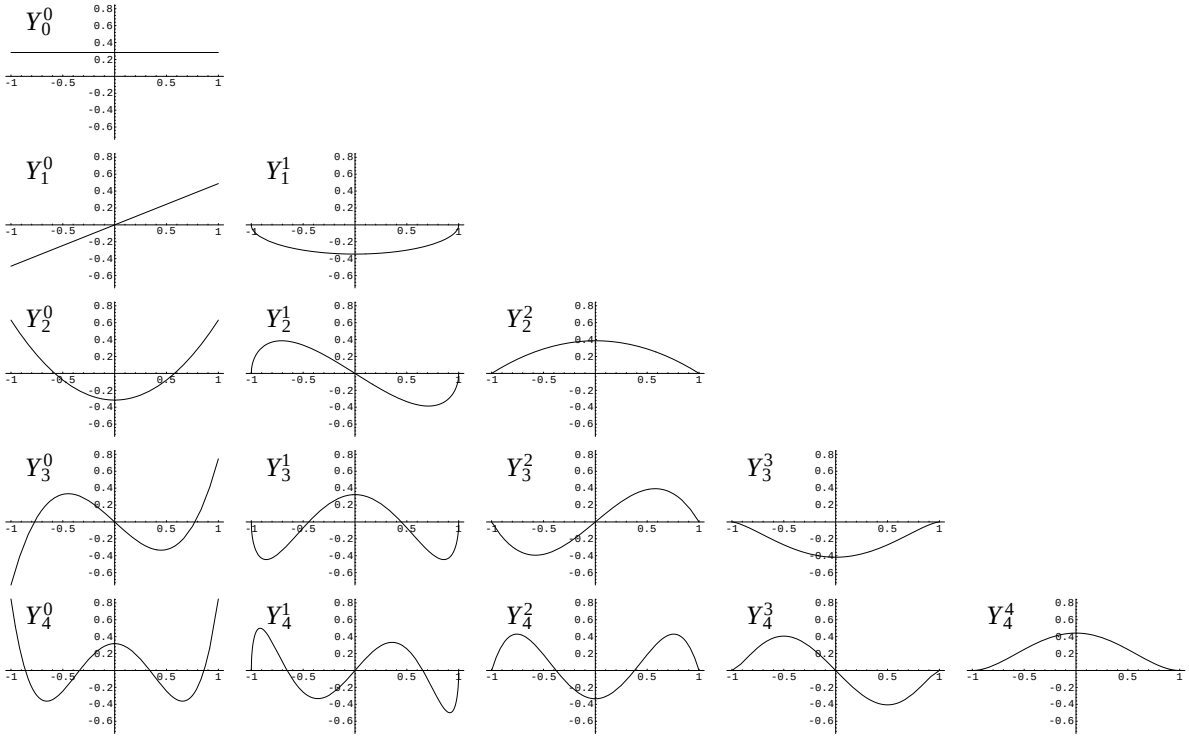


Figure 5.1: Plots of $Y_l^m(\theta = \arccos w, \phi = 0)$ up to degree 4. These diagrams are scaled plots of the associated Legendre polynomials.

Figure 5.1 plots $Y_l^m(\arccos w, 0)$ for $0 \le m \le l < 5$. The functions Y_l^m take real values for $\phi = 0$. The graphs give a quantitative impression of the zonal amplitude variation of Y_l^m. At the same time, they are plots of the associated Legendre polynomials, except for a scaling constant (cf. (5.3)). The leftmost column corresponds to the Legendre polynomials themselves. Figure 5.2 gives a more qualitative impression of our basis functions and their signs. The real parts of the spherical harmonics up to degree 5 are displayed as gray levels on the surface of the unit sphere,

l	$m=0$	$m=1$	$m=2$	$m=3$
0	1			
1	$\cos\theta$	$e^{i\phi}\,\sin\theta$		polar
2	$-1+3\cos^2\theta$	$e^{i\phi}\,\cos\theta\,\sin\theta$	$e^{2i\phi}\sin^2\theta$	
3	$-3\cos\theta+5\cos^3\theta$	$e^{i\phi}(1-5\cos^2\theta)\sin\theta$	$e^{2i\phi}\cos\theta\,\sin^2\theta$	$e^{3i\phi}\sin^3\theta$
0	1			
1	u_2	$u_0+i\,u_1$		Cartesian
2	$-1+3\,{u_2}^2$	$(u_0+i\,u_1)\,u_2$	$(u_0+i\,u_1)^2$	
3	$-3\,u_2+5\,{u_2}^3$	$(u_0+i\,u_1)(1-5{u_2}^2)$	$(u_0+i\,u_1)^2u_2$	$(u_0+i\,u_1)^3$
0	1			
1	u_2	u_0		Re Y
2	$-1+3\,{u_2}^2$	$u_0\,u_2$	${u_0}^2-{u_1}^2$	
3	$-3\,u_2+5\,{u_2}^3$	$u_0\,(1-5\,{u_2}^2)$	$({u_0}^2-{u_1}^2)u_2$	$u_0^3-3u_0u_1^2$
0	0			
1	0	u_1		Im Y
2	0	$u_1\,u_2$	$u_0\,u_1$	
3	0	$u_1(1-5{u_2}^2)$	$u_0\,u_1\,u_2$	$u_1^3-3u_0^2u_1$
0	$1/\sqrt{4\,\pi}$			
1	$\sqrt{3/4\,\pi}$	$\sqrt{3/8\,\pi}$		
2	$\sqrt{5/16\,\pi}$	$\sqrt{15/8\,\pi}$	$\sqrt{15/32\,\pi}$	
3	$\sqrt{7/16\,\pi}$	$\sqrt{21/64\,\pi}$	$\sqrt{105/32\,\pi}$	$\sqrt{35/64\,\pi}$

Table 5.1: Explicit expressions of the spherical harmonics up to degree 3, in both polar and Cartesian form. The last part of the table gives the common normalizing constants, e.g., $Y_1^0=\sqrt{3/4\pi}\,u_2$.

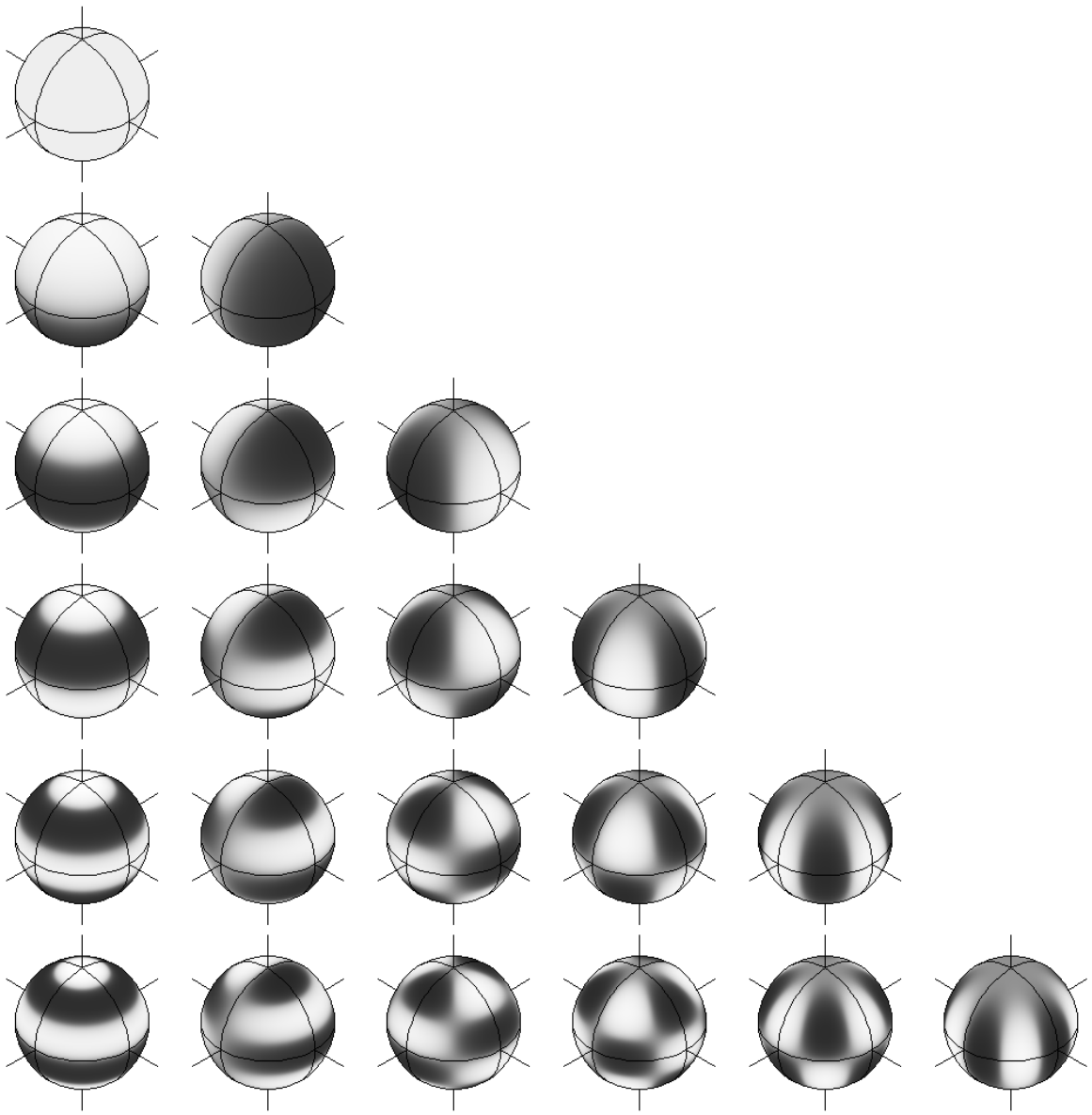


Figure 5.2: The real parts of the spherical harmonic functions Y_l^m, with l growing from 0 (top) to 5 (bottom), and m ranging from 0 (left) to l in each row. The function value is mapped on the spheres, light grey representing positive values and dark grey negative values.

which is the domain of Y_l^m. The real-valued functions Y_l^0 appear by themselves in the left column (m =0). For $m > 0$, the imaginary part is the same as the real part, rotated by $-\frac{\pi}{2m}$.

A spherical harmonic function of degree l is a polynomial of degree l in u_0, u_1 and u_2. It can be written as a homogeneous polynomial of degree l (using the identity ${u_0}^2 + {u_1}^2 + {u_2}^2 = 1$, on Ω_3).

In some cases, we have to adopt an indexing scheme $j(l,m)$ that assigns a unique index j to every pair l, m, like e.g. $j(l,m) := l^2 + l + m$. When the degree of the spherical harmonics is limited, i.e. $0 \le l < n_l$, j is also limited by $j < n_j = n_l^2$.

5.1.2 Expressing surface shape

With spherical harmonics the series (2.3) takes the form

$$\underline{x}(\theta,\phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} \underline{c}_l^m \; Y_l^m(\theta,\phi) \; . \tag{5.5}$$

The coefficients

$$\underline{c}_l^m = \begin{pmatrix} c_{l\,0}^m \\ c_{l\,1}^m \\ c_{l\,2}^m \end{pmatrix}$$

in this series are three-dimensional vectors. Their components, $c_{l\,0}^m$, $c_{l\,1}^m$ and $c_{l\,2}^m$, are complex numbers for $m \neq 0$ in general; they are real numbers for $m = 0$.

For convenience, the real and imaginary parts of the complex basis functions Y_l^m can be used as independent real valued basis functions. The set of functions $\{Y_l^0, \mathrm{Re}(Y_l^m), \mathrm{Im}(Y_l^m)\}$ (where $l \ge 0$ and $0 < m \le l$) is orthogonal but not normalized. For $m > 0$, $\oint \mathrm{Re}(Y_l^m)^2 = \oint \mathrm{Im}(Y_l^m)^2 = \frac{1}{2}$. The real functions might be scaled by $\sqrt{2}$ for $m \neq 0$. There is the same number of functions as with the complex basis, namely $2l+1$ for any non-negative l, or n_l^2 in total.

5.1.3 Variability of spherical harmonic surfaces

All of the following shapes are defined as spherical harmonic surfaces of degree up to three, i.e. $n_l = 4$. Each of x_0, x_1 and x_2 are defined by the 16 coefficients in the corresponding series. They illustrate the wide variability that can be achieved even with a low degree. Typically only

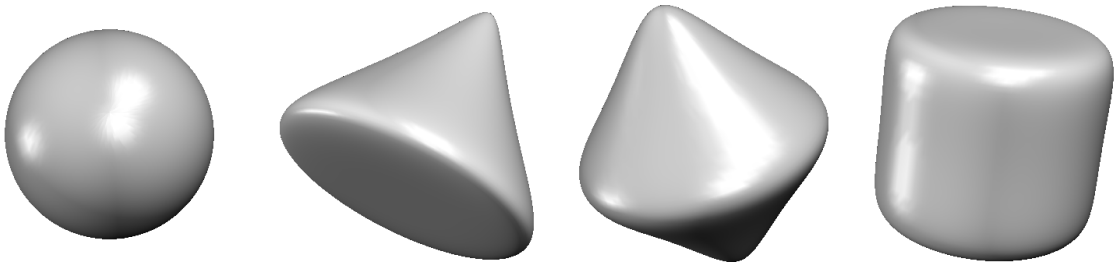


Figure 5.3: Smooth rotational geometrical objects: vector valued series of spherical harmonics represent a sphere, and they approximate a cone, a double cone, and a cylinder.

a few (like 5 or 9) of the 48 coefficients are different from zero in the examples.

The basis functions are smooth. The objects described by a truncated series of the form (5.5) tend to be smooth as well (Fig. 5.3). But they can have sharp edges or cusps as well, as Figure 5.4 illustrates. The first three

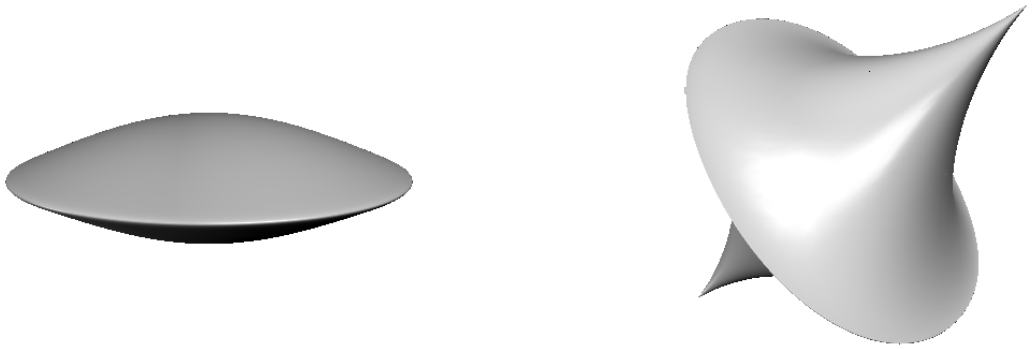


Figure 5.4: The lens has high curvature at the rim, and the top has a sharp edge and points. Yet they are composed of low degree spherical harmonics.

platonic polyhedra (Figure 5.5) can be modelled with the truncated set of basis functions. The polyhedra in Figure 5.5 have rotational symmetries of order two and three, and the cube and the octahedron have even fourth order symmetry. More generally, an object using spherical harmonics up to degree $n_l - 1$ can have at most n_l-ary rotational symmetry. Examples for $n_l = 2 \ldots 4$ appear in Figure 5.6.

There is no reason to limit ourselves to star-shaped objects. Figure 5.7 lists a few that are not. The Dodecahedron and the Icosahedron would need higher harmonics: but the tetrahedron uses only degree one and

Figure 5.5: The tetrahedron, the cube or hexahedron, and the octahedron are the three simpler ones of the five platonic polyhedra.

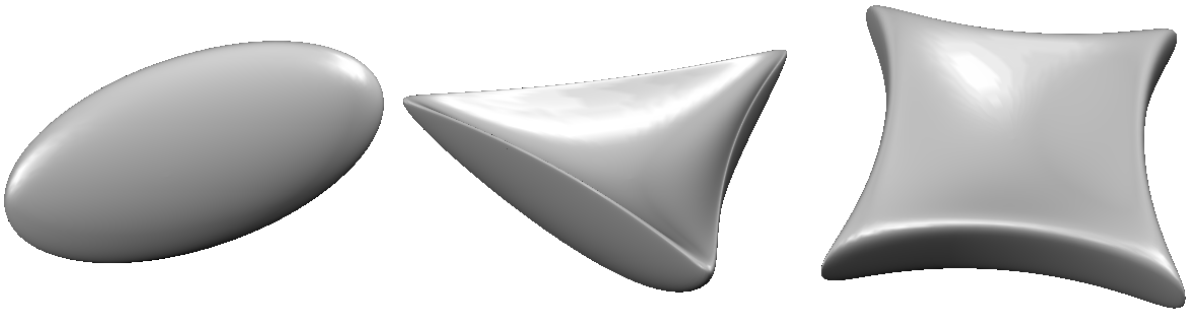


Figure 5.6: Objects involving spherical harmonics up to degree 1, 2 and 3 can have a 2-, 3- and 4-fold rotational symmetry.

two. The cube and the octahedron have no second degree component. They only differ in the sign of the third degree contributions: starting from a sphere, the points that move out on the cube move in on the octahedron, and vice versa. The upper signs in (5.6) correspond to the cube, the lower ones to the octahedron.

$$\begin{pmatrix} x_0 \\ x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -0.7071\ \mathrm{Re}(Y_1^1) \pm 0.0745\ \mathrm{Re}(Y_3^1) \pm 0.098\ \mathrm{Im}(Y_3^3) \\ -0.7071\ \mathrm{Im}(Y_1^1) \pm 0.0745\ \mathrm{Im}(Y_3^1) \pm 0.098\ \mathrm{Re}(Y_3^3) \\ 0.5\ Y_1^0 \mp 0.0863\ Y_3^0 \end{pmatrix} \quad (5.6)$$

Objects with symmetries have been used in the examples up to now because their whole shape can be perceived or at least guessed from a single view. The chosen objects are special cases with respect to symmetry and not representative in that sense for shapes expressable with spherical harmonic descriptors. Figure 5.8 gives an example of a general object.

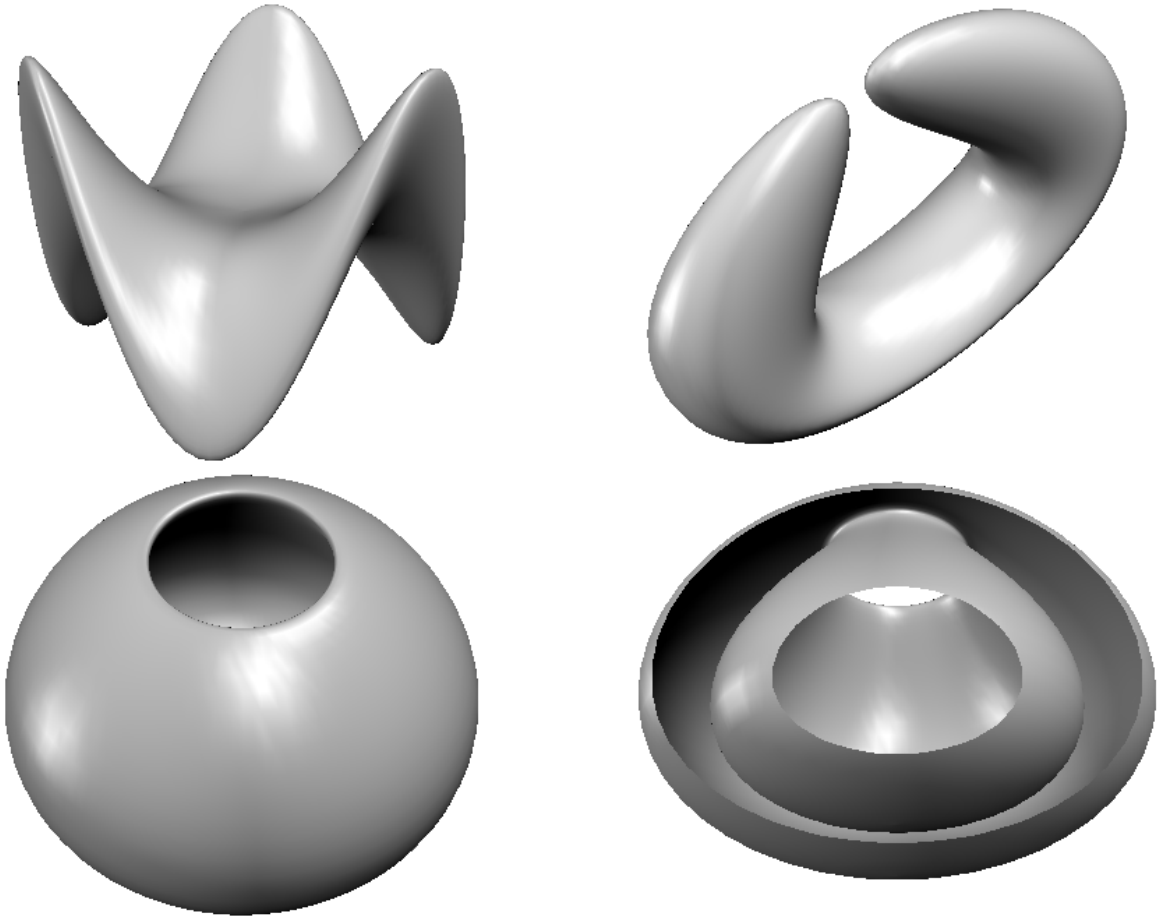


Figure 5.7: Neither of these is star-shaped: A monkey saddle, a croissant and a vase. In the last image, the vase is cut open to reveal the folding in of the surface.

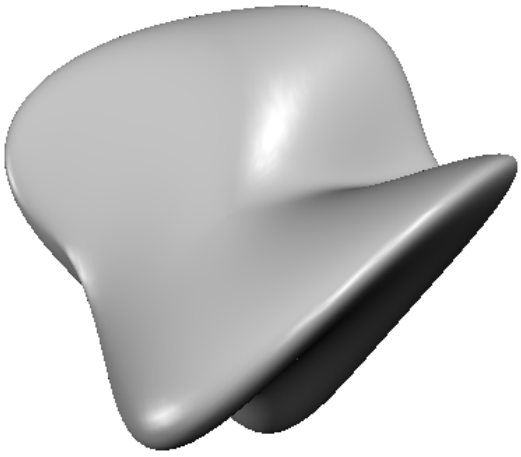
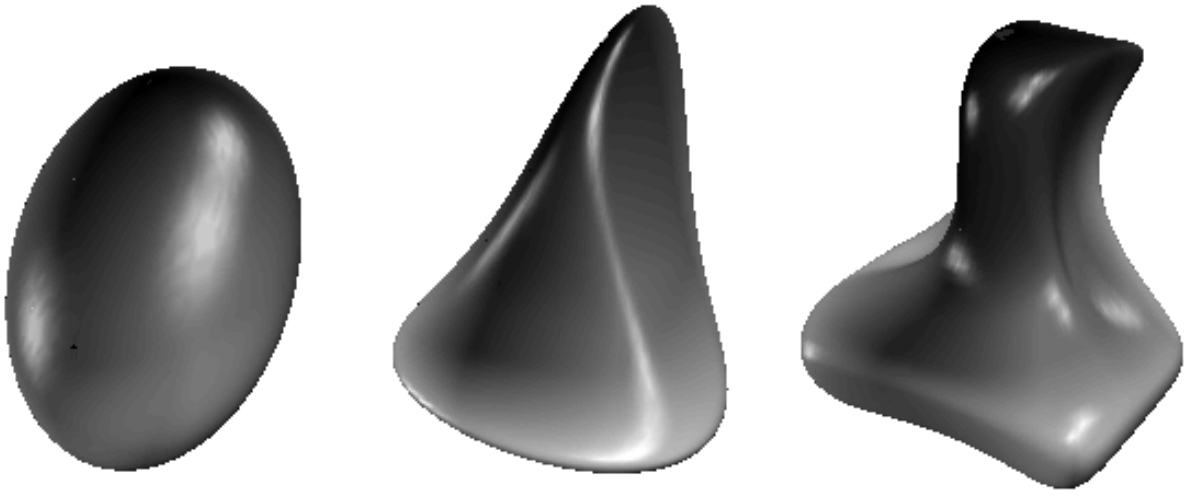


Figure 5.8: A "random" asymmetric blob shaped object

5.2 Harmonic shape descriptors

Section 3.4 introduced the parametrization of the surface of a simply connected object. The parametrization provides a correspondence between any surface vertex x_i and its parameter u_i, which can be interpreted as the sampling of a function $\underline{x}(\underline{u})$, i.e. $\underline{u}[i] \rightarrow \underline{x}(\underline{u}[i]) = \underline{x}[i]$. Now $\underline{x}(\underline{u})$ is expanded into a series of spherical harmonics as in (5.5), or into any other set of basis functions

The coefficients of the spherical harmonic functions of different degrees provide a measure of the spatial frequency constituents of the structure. Partial sums of the series (5.5) for the "duck" test object are plotted in Figure 5.9. The sums are truncated by limiting l to $0 \leq l < n_l$, where $n_l = 2, 4, 8$. As higher frequency components are included, more detailed features of the object appear.



(a) (b) (c)

Figure 5.9: Global shape description by expansion into spherical harmonics: The figures illustrate the reconstruction of the partial spherical harmonic series, using coefficients up to degree 1 (a), to degree 3 (b) and 7 (c).

5.2.1 Integration over the sphere

The use of orthonormal basis functions is convenient for the calculation of the expansion coefficients. Formally, the coefficients are calculated by forming the inner product of $\underline{x}$ with the basis function in question:

$$\underline{c}_l^m \quad = \quad \langle \underline{x}(\theta,\phi), Y_l^m(\theta,\phi) \rangle = \int_0^{\pi} \int_0^{2\pi} \underline{x}(\theta,\phi)\, Y_l^m(\theta,\phi)\, d\phi \sin\theta\, d\theta \qquad (5.7)$$

The parametrization defines the function $\underline{x}(\theta,\phi)$ only for the parameter coordinates of the vertices. Only $\underline{x}(\theta_i,\phi_i) = \underline{x}_i$ is defined, where i is the index of a vertex, $0 \leq i < n_{\text{vert}}$. For the evaluation of the integral (5.7) we would have to define an interpolating function between these sample points; an adaptation of bilinear interpolation could be used for this purpose. But this would introduce an artificial sub-voxel resolution that is not supported by the input data. On the other hand, the straightforward discretization of the integral is

$$\underline{c}_l^m \quad \approx \quad \sum_{i=0}^{n_{\text{vert}}-1} \underline{x}_i\, Y_l^m(\theta_i,\phi_i)\, \Delta\Omega \qquad , \qquad (5.8)$$

with the finite $\Delta\Omega$ replacing $d\Omega = d\phi \sin\theta d\theta$. There are two possibilities for choosing $\Delta\Omega$. Setting $\Delta\Omega = \frac{4\pi}{n_{\text{vert}}}$ distributes the area of Ω_3 evenly over all vertices. But it might be argued that the area of a facets is relevant, and hence should be distributed to its four corners. Then $\Delta\Omega = \frac{4\,\pi\, count_i}{4\, n_{\text{face}}}$, where $count_i$ is the number of neighbors of vertex i. But neither of these schemes in general gives the precise coefficients of a series representing our object. The reason is that although the functions Y_l^m are orthonormal, their values evaluated at some set of parameter pairs (θ_i,ϕ_i) will generally not form an orthonormal set of vectors.

We can arrange all needed values of our basis functions in a $n_{\text{vert}} \times n_j$ matrix B where $b_{i,j(l,m)} = Y_l^m(\theta_i,\phi_i)$. In the usual case where n_j is significantly smaller than n_{vert}, the columns of B are approximately orthogonal. We further arrange the object space coordinates of all vertices in an $n_{\text{vert}} \times 3$ matrix $X = (\underline{x}_0, \underline{x}_1, \ldots \underline{x}_{n_{\text{vert}}-1})^T$ and all coefficients in the $n_j \times 3$ matrix $C = (\underline{c}_0^0, \underline{c}_1^{-1}, \underline{c}_1^0 \ldots)^T$. The equations (5.8) for all l and m take the compact form $C \approx \frac{4\pi}{n_{\text{vert}}} B^T X$. But what we really want is a spherical harmonic series that passes near the real positions of our vertices, i.e. $X = BC + E$ where the error matrix E should be small. These so-called normal equations are solved with least square sums over the columns of E by

$$C \quad = \quad (B^T B)^{-1} B^T X \quad . \qquad (5.9)$$

The formula 5.9 mathematically states the use of the pseudo-inverse of B: it does not imply a numerical evaluation of the matrix expressions on

the right. The global approximation error is the square of the Frobenius norm of $E = BC - X$, which is also minimized. This is not too different from (5.8) because the symmetric $n_j \times n_j$ matrix $\frac{4\pi}{n_{\text{vert}}} B^T B$ is close to the identity matrix.

5.2.2 Invariant descriptors

The coefficients obtained thus far still depend on the relative position of the parameter net of the object surface and on the orientation of the object in space (Figure 3.6). We can get rid of these dependencies by rotating the object to canonical positions in parameter space and object space. This needs three rotations in parameter space and three rotations in object space, when rotations are described using Euler angles. All rotations result in new linear combinations of the components of the harmonic descriptors.

The relations between the Cartesian and the spherical coordinates of the parameter space are $u_0 = \sin\theta \cos\phi$, $u_1 = \sin\theta \sin\phi$ and $u_2 = \cos\theta$. To define a standard position we consider only the contribution of the spherical harmonics of degree $l = 1$ in equation (5.5).

$$\underline{x}_1(\theta, \phi) = \sum_{m=-1}^{1} \underline{c}_1^m \; Y_1^m(\theta, \phi) \tag{5.10}$$

This sum involves the basis functions $Y_1^{-1} = \frac{\sqrt{3}}{2\sqrt{2\pi}} (u_0 - iu_1)$, $Y_1^0 = \frac{\sqrt{3}}{2\sqrt{\pi}} u_2$ and $Y_1^1 = -\frac{\sqrt{3}}{2\sqrt{2\pi}} (u_0 + iu_1)$. Any three real valued linear combinations [1] of these, interpreted as Cartesian coordinates in the object space, will always describe an ellipsoid(see Figure 5.9 a). We rotate the object in parameter space so that the north pole ($\theta = 0$, on the u_2 axis) will be at one end of the shortest main axis of this first order ellipsoid and the point where the Greenwich meridian ($\phi = 0$) crosses the equator ($\theta = \frac{\pi}{2}$, on the u_0 axis) is at one end of the longest main axis.

This paragraph explains how I determine the main axes. At the three main axes, the length of the vector $\underline{x}_1(\theta, \phi)$ is stationary: it reaches a maximum, a saddle point, and a minimum, respectively. Measuring Euclidean lengths becomes simpler when we transform the component vectors to a Euclidean, real valued form. Applying the definitions of the

[1] The combination $\underline{x}_1$ is real if and only if $\left(\underline{c}_1^{-1}\right)^* = -\underline{c}_1^1$ and $\underline{c}_1^0 \in \mathbb{R}^3$

Y_1^m yields

$$\underline{x}_1(\underline{u}) \;=\; A\underline{u} \;=\; A\begin{pmatrix} u_0 \\ u_1 \\ u_2 \end{pmatrix} \;=\; \underline{a}_1\, u_0 + \underline{a}_2\, u_1 + \underline{a}_3\, u_2 \quad , \tag{5.11}$$

where

$$A = (\underline{a}_1, \underline{a}_2, \underline{a}_3) \;=\; \frac{\sqrt{3}}{2\sqrt{2\pi}}\left(\underline{c}_1^{-1} - \underline{c}_1^1, i(\underline{c}_1^{-1} + \underline{c}_1^1), \sqrt{2}c_1^0\right) \quad . \tag{5.12}$$

We are looking for the unit vectors $\hat{\underline{u}}_1$, $\hat{\underline{u}}_2$ and $\hat{\underline{u}}_3$ that maximize or minimize the length of the vector. The solutions are the eigenvectors of $A^T A$, with nonnegative eigenvalues $l_1^2 > l_2^2 > l_3^2$. Their roots l_1, l_2 and l_3 represent half the lengths of the main axes of the ellipsoid. At the middle eigenvector, $\hat{\underline{u}}_2$, $\|\underline{x}_1\|$ has a saddle-point rather than an extremum. The normalized eigenvectors form the rotation matrix $R_u^T = (\hat{\underline{u}}_1, \hat{\underline{u}}_2, \hat{\underline{u}}_3)$, which is applied to the parameters $\underline{u}[i]$ associated with each vertex i: $\underline{u}[i]' = R_u^T\, u_i$. This new parametrization results in new coefficients $\underline{c}_l^{m'}$ and hence in the new coefficient matrix $A' = A\, R_u$. Its three column vectors $\underline{a'}_1$, $\underline{a'}_2$ and $\underline{a'}_3$ are the main axes of the first order ellipsoid in object space.

All rotations are determined based on the values of $\underline{c}_1^{m'}$ $(m' = -1 \ldots 1)$ of the ellipsoid only, but they are applied to all components of the descriptor $\{\underline{c}_l^m\}$. The parameter space rotations result in a different description of the same object in the same position, just parametrized in a standard way.

Now, the ellipsoid is rotated in the object space to make its main axes coincide with the coordinate axes, putting the longest ellipsoid axis along x_0 and the shortest one along x_2. The object space rotations require only the matrix multiplication $\underline{c}_l^{m''} = R_x\, \underline{c}_l^{m'}$. The object space rotation matrix is $R_x = \text{diagonal}(\frac{1}{l_1}, \frac{1}{l_2}, \frac{1}{l_3}) \cdot A'^T$. It rotates the main axes of the ellipsoid into an axis-parallel position and makes the coefficient matrix $A'' = R_x\, A' = R_x\, A\, R_u$ diagonal. The elements of the diagonal are the lengths of the main axes of the ellipsoid.

Parameter Visualization

This subsection introduces a specific surface pattern. The pattern is gray valued and covers the whole surface of an object. It always stays fixed with relation to the parameter space Ω_3. This makes it possible to

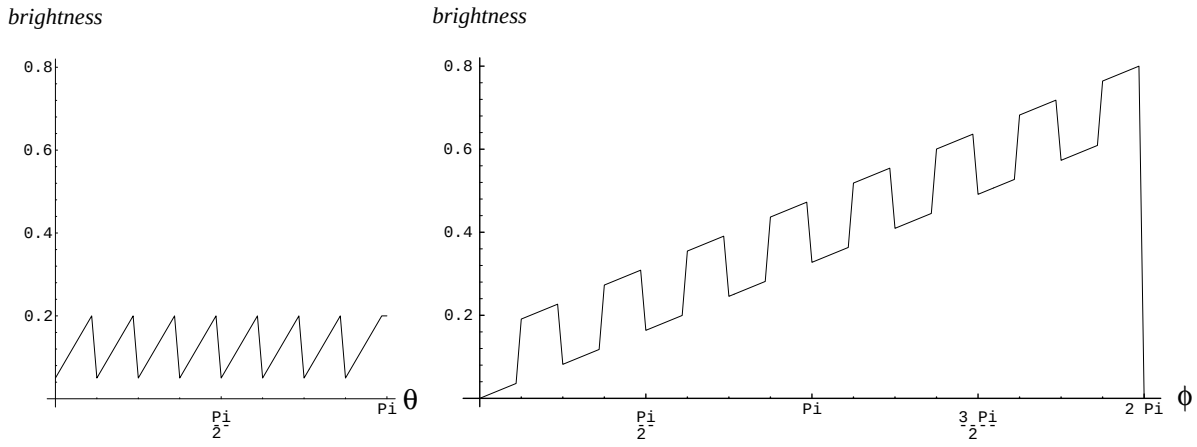


Figure 5.10: The components of the pattern are a sawtooth function of θ (*left*), and a ramp plus a rectangle function of ϕ (*right*).

estimate the parameter value associated with a location on the surface from the shading of is neighborhood. The pattern makes visible the parameter space rotations described above.

The brightness of the pattern varies with θ in a sawtooth fashion, cf. Figure 5.10a, and with ϕ in a ramp and rectangle function, cf. Figure 5.10b. The whole pattern is the additive superposition of the θ and ϕ contributions; Figure 5.11 presents it. The diagrams of Figure 5.10

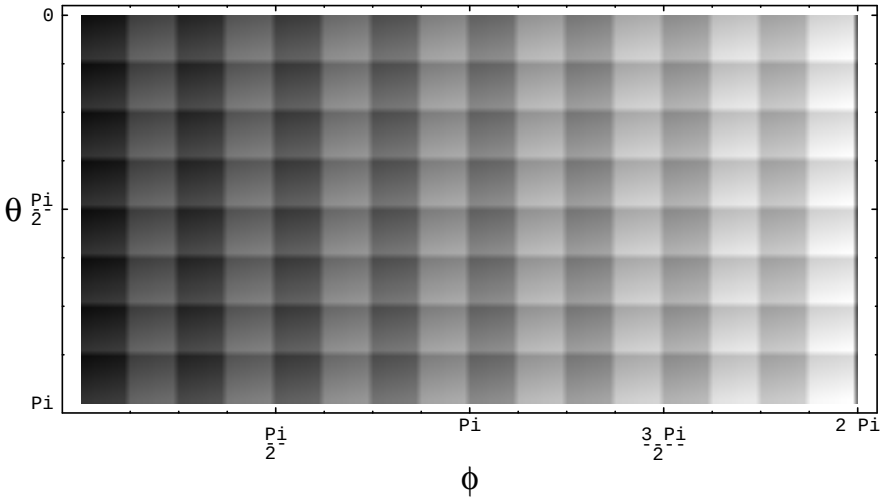
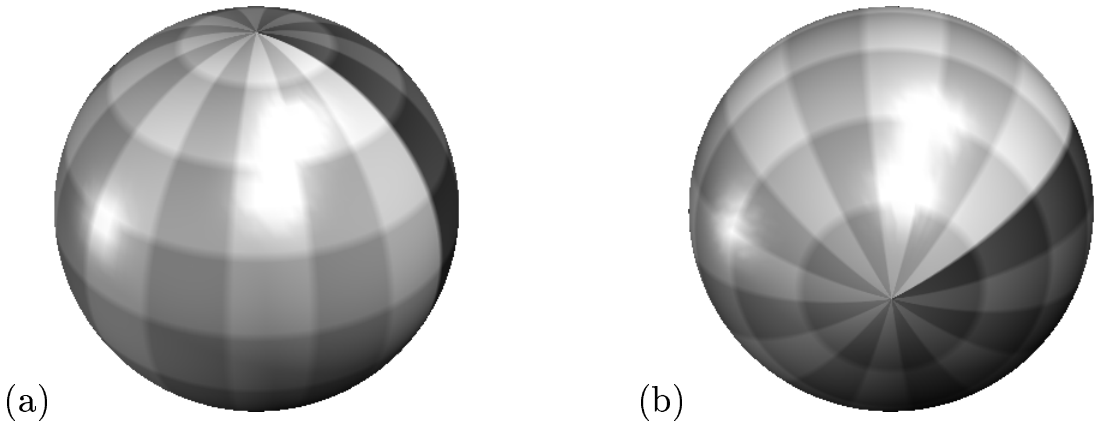


Figure 5.11: This pattern, when mapped to an object surface, reveals the pose of the parameter space. The θ axis points down to preserve a right handed coordinate system. The top border $\theta = 0$ corresponds to the north pole, the bottom border $\theta = \pi$ is the south pole.

illustrate that the functions interpolate linearly between sample points spaced $\frac{\pi}{64}$ apart and that there are no discontinuities (jumps). Even the descent in the wrap around for ϕ extends from $\frac{127\pi}{64}$ to $2\pi \equiv 0$ and is continuous.



(a) (b)

Figure 5.12: The simplest object, a sphere, exemplifies the interpretation of the surface pattern. The specular highlights don't belong to the pattern but are added for a better 3D impression. *(a)* The north pole is visible from the standard view point. We perceive it as the center of a bright circular zone in spite of its dark center (the pole itself). *(b)* The region of the south pole, on the contrary, appears as a dark zone, although the pole has actually a light shade.

In Figure 5.12 the surface pattern covers the sphere $\underline{x} = \underline{u}$. This is the simplest possible object, because here only first degree terms contribute, and the object space coordinates are identical with the parameter space coordinates. Figure 5.12 might be called a "picture of the parameter space". It should help the interpretation of the following pictures that use the same pattern for parameter coding. The sphere is striped lengthwise with 16 sectors, each 8 lighter and darker ones. The stripes meet at the poles. The shading becomes slowly brighter as ϕ increases, i.e. towards east (cf Fig. 5.10b). With the normal (right handed) orientation of the parameter space, this gives a positive or counterclockwise increase of the brightness around the north pole and a negative or clockwise increase around the south pole. After one turn, the intensity leaps back from brightest to reach the darkest value on the Greenwich meridian. When moving southward, i.e. with increasing θ, the shading gradually becomes lighter, only to drop back to the same dark level every $\frac{\pi}{8}$ (cf Fig. 5.10a). The slow brightening from the north pole to the first maximum at $\pi/8$ is less visible than the rapid drop that follows, so that the north pole

appears as the center of a bright disk (but this center is dark!). The opposite is true at the south pole, which is the bright center of circle with a dark border[2].

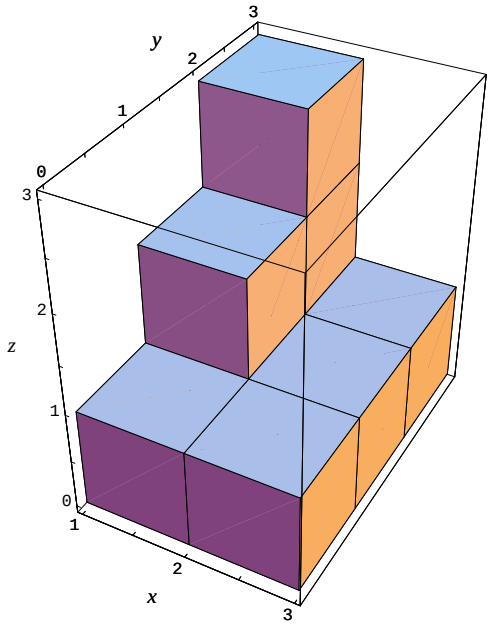


Figure 5.13: This nine voxel object will illustrate the standardization of descriptors and their comparison, including symmetry considerations. The object itself is completely asymmetric. Two steps of a stair sit on top of a 3 by 2 base plate. We call the object "stair".

To illustrate the descriptor and its rotations in parameter and object space, a test object should be very simple, but it must not have symmetries with respect to any plane, straight line, or point. Symmetries are frequent in few-voxel objects, but they would make the pose ambiguous and interfere with the symmetry discussions below in subsection 5.3.1. The nine voxel object in Figure 5.13 is fit for the purpose.

Parametrization of the surface and expansion into the spherical harmonic basis yields a descriptor for the object. Partial series of (5.5) define surfaces of increasing levels of detail. They visualize the descriptor in Figure 5.14, and they illustrate its hierarchical organization from coarse to fine. Thanks to the surface pattern, the relation to the parameter space Ω_3 is visible.

Figure 5.15 shows the descriptor $\underline{c}$ before any standardization. The

[2]The intensity drop that should happen exactly at the south pole is suppressed. The isolated dark spot would be distracting.

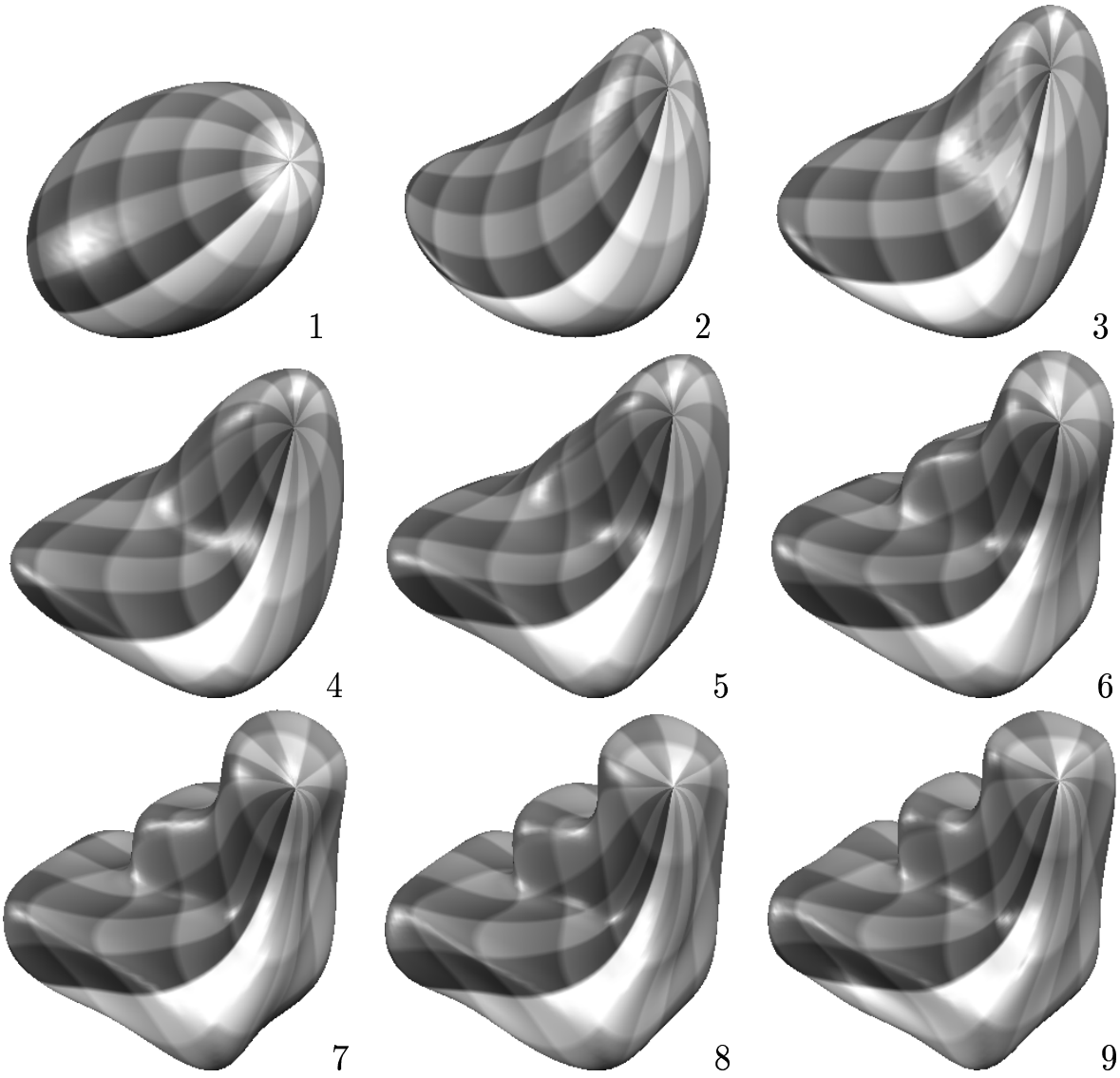


Figure 5.14: The "stair" object, reconstructed from its descriptor. The numbers indicate the maximum degree l in the partial sum, i.e. $n_l - 1$. As higher frequency components are added, more and more details show up. This is not the standard view: x_0 increases towards front and left, x_1 towards front and right, x_2 upwards. The new view point gives a better look at the north pole.

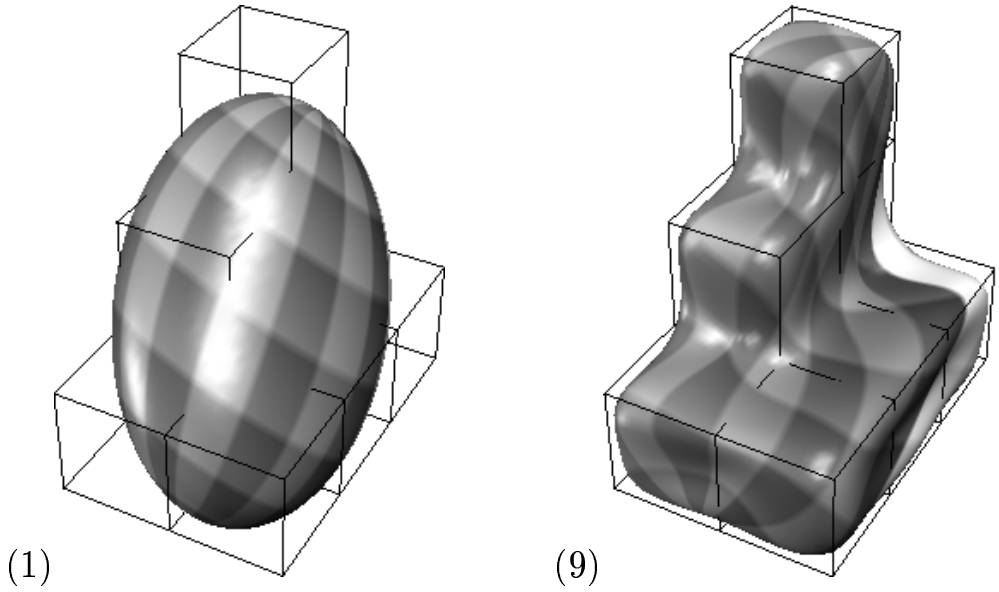


Figure 5.15: The raw descriptor of the "stair" object. *Left:* The first degree ellipsoid. *Right:* Reconstruction up to degree 9 ($n_l = 10$). The edges of the original object give an impression of the accuracy of the partial sums. Standard view: x_0 increases towards the right and front, x_1 to the back and right, and x_2 increases upwards.

subsequent steps take as their reference the first degree ellipsoid, which appears in the left image. Their effects on both the ellipsoid and the full object ($n_l = 10$) are shown in each case.

Parameter space rotation takes us to Figure 5.16, showing $\underline{c}'$. The poles ($\theta = 0$ and $\theta = \pi$) end up on the shortest main axis, and the Greenwich meridian ($\phi = 0$) passes through the longest main axis. The resulting descriptor still represents exactly the same geometrical surface. The superimposed wireframe of the original "stair" object confirms that the object has not moved in object space. The object space rotation now leads to a descriptor $\underline{c}''$ in canonical position (Figure 5.17). The edges of the cube $\left[0, \sqrt{3/8\pi}\right]^3$ are overlaid as a reference of size and orientation. The main axes of the ellipsoid and of its parametrization line up with the coordinate axes. The final scaling makes the coefficient of $\mathrm{Re}(Y_1^1)$, which corresponds to the longest main axis, equal to 1. The half length of the longest axis becomes $-Y_1^1(\frac{\pi}{2}, 0) = -f_1^1((1,0,0)^T) = \sqrt{3/8\pi}$. For the descriptor of a reasonably sized object, this is a shrinking by several orders of magnitude.

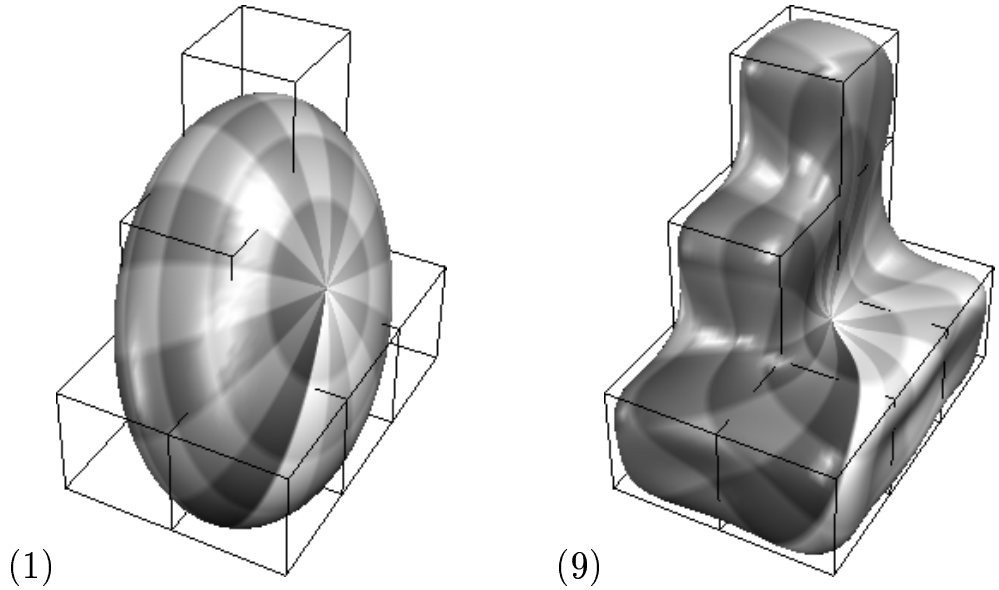


Figure 5.16: The descriptor of the "stair" object after parameter space rotation. *Left:* The first degree ellipsoid. *Right:* Reconstruction up to degree 9 ($n_l = 10$). Standard view. The wireframe of the original object is overlaid again.

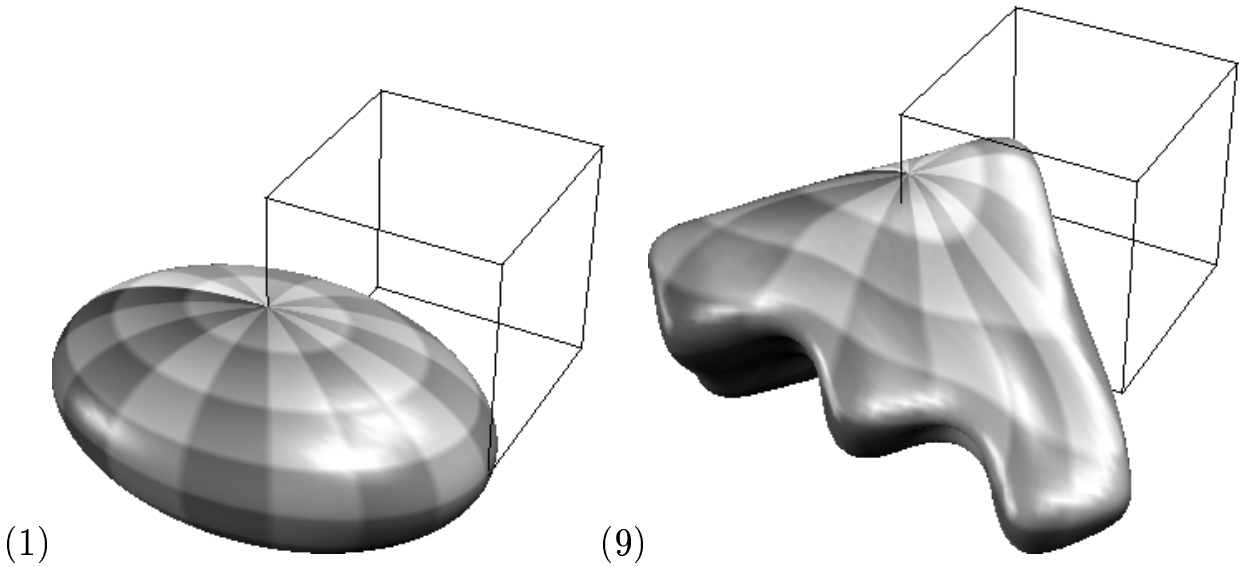


Figure 5.17: The descriptor of the "stair" object in canonical position. *Left:* The first degree ellipsoid. *Right:* Reconstruction up to degree 9 ($n_l = 10$).

The descriptors $\underline{c}''$ are now invariant under rotation of the object, except mirrorings (rotations by π). Including information from higher degree coefficients could eventually disambiguate these cases. Subsection 5.3.1 presents an alternative approach, which considers all possible mirrorings. Ignoring $\underline{c}_0^0$ results in translation invariance. Scaling invariance can be achieved by dividing all descriptors by l_1, the length of the longest main axis.

5.2.3 Importance of uniform parametrization

This thesis thus far assumed that a homogeneous density and a minimal distortion of the parameter net would be important for shape characterization, especially for obtaining an invariant description. Similarly, the 2-D expansion of contours $s(t)$ into series of harmonics [23] was based on the model of tracing a curve with constant velocity, i.e. assigning same lengths ΔL to equivalent parameter steps Δt. A non-uniform distribution of parameters on an object surface, e.g. by clustering at certain locations, seems to be suboptimal with respect to a uniform representation of the whole surface. One would expect an over-representation of some parts at the expense of other regions, resulting in a distorted shape description.

The importance of a parametrization with minimal distortion can be demonstrated with an experiment. The expansion into a series of spherical harmonics is calculated for both the non-uniform initial parametrization (bypassing the optimization step for this part of the experiment) and the result after optimization. A manifestly non-star-shaped form was chosen: the original object consists of 11 voxels and is shaped like the character E. Its initial and optimized parametrizations are given in Figure 5.18 for comparison. The diagrams correspond to Figures 4.2a and 3.6a.

Figure 5.19 illustrates the expansion in a spherical harmonic series up to degree ten and the truncated reconstruction up to degree one (top), four (middle) and ten (bottom) for the initial (left) and optimized (right) parametrization. (A five-fold oversampling was applied to the surface to represent it accurately. This may be viewed as a rough form of numerical integration.) Comparing the expressive difference, one can conclude that a uniform parametrization is absolutely essential to obtain useful spherical harmonic descriptors. Even from the distorted initial parametrization, descriptors can be derived, that are necessarily "optimal" in the least squares sense, but the series of harmonics does not reflect the shape properties of the surface. Using the optimized parametrization, coeffi-

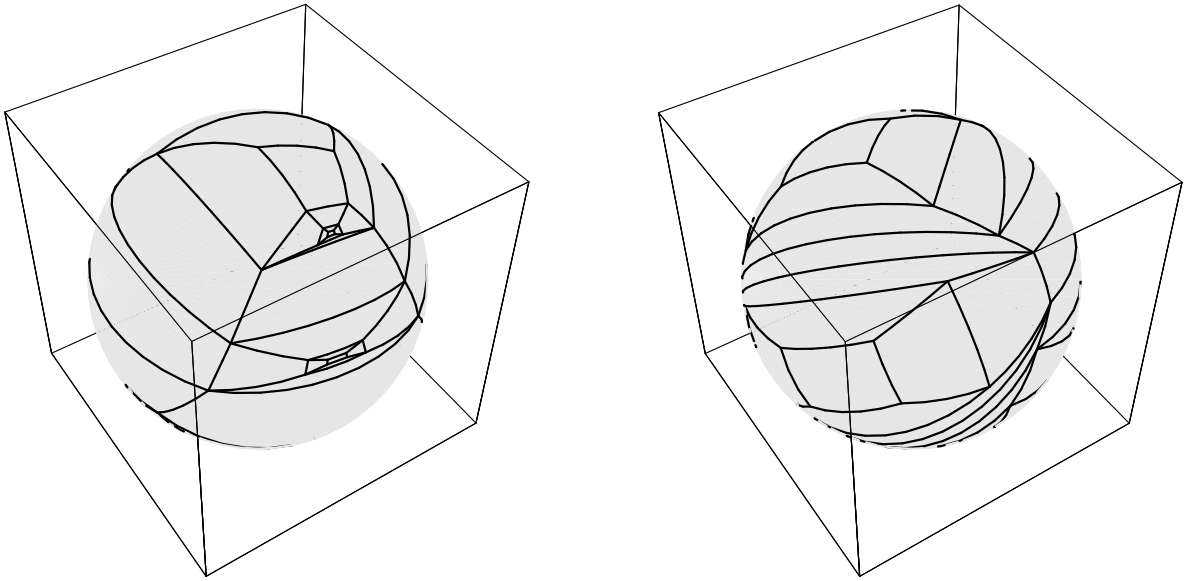


Figure 5.18: Two different parametrizations of the "E" object. *Left:* The initial parametrization is the starting point of the optimization. *Right:* The optimized parametrization.

cients of higher degree add information about details of higher spatial frequency; this is desirable. The first degree harmonic approximation (Figure 5.19b) covers the whole object and comprises information about the major size and elongation, whereas the three "legs" of the E-shape appear in the reconstruction using harmonics up to degree four (Figure 5.19d,f).

5.3 Comparison and Recognition of Shapes

The main purpose of invariant spherical harmonic descriptors is to transform the task of comparing the shape of two objects to the simpler comparison of their descriptors.

In the truncated series, l takes the n_l different values $0, 1, \ldots n_l - 1$. The descriptor presents itself as n_l^2 three-vectors or as one flat $3\,n_l^2$-vector. Thanks to the invariance properties of the descriptors, the similarity between the feature vectors of two objects measures the similarity between their shapes. One of the simplest dissimilarity measures is the Euclidean distance between the feature vectors. This is the same as the L_2 norm of the difference of the two vectors. The square of this quantity is the sum of the squared differences between corresponding entries in the two vectors.

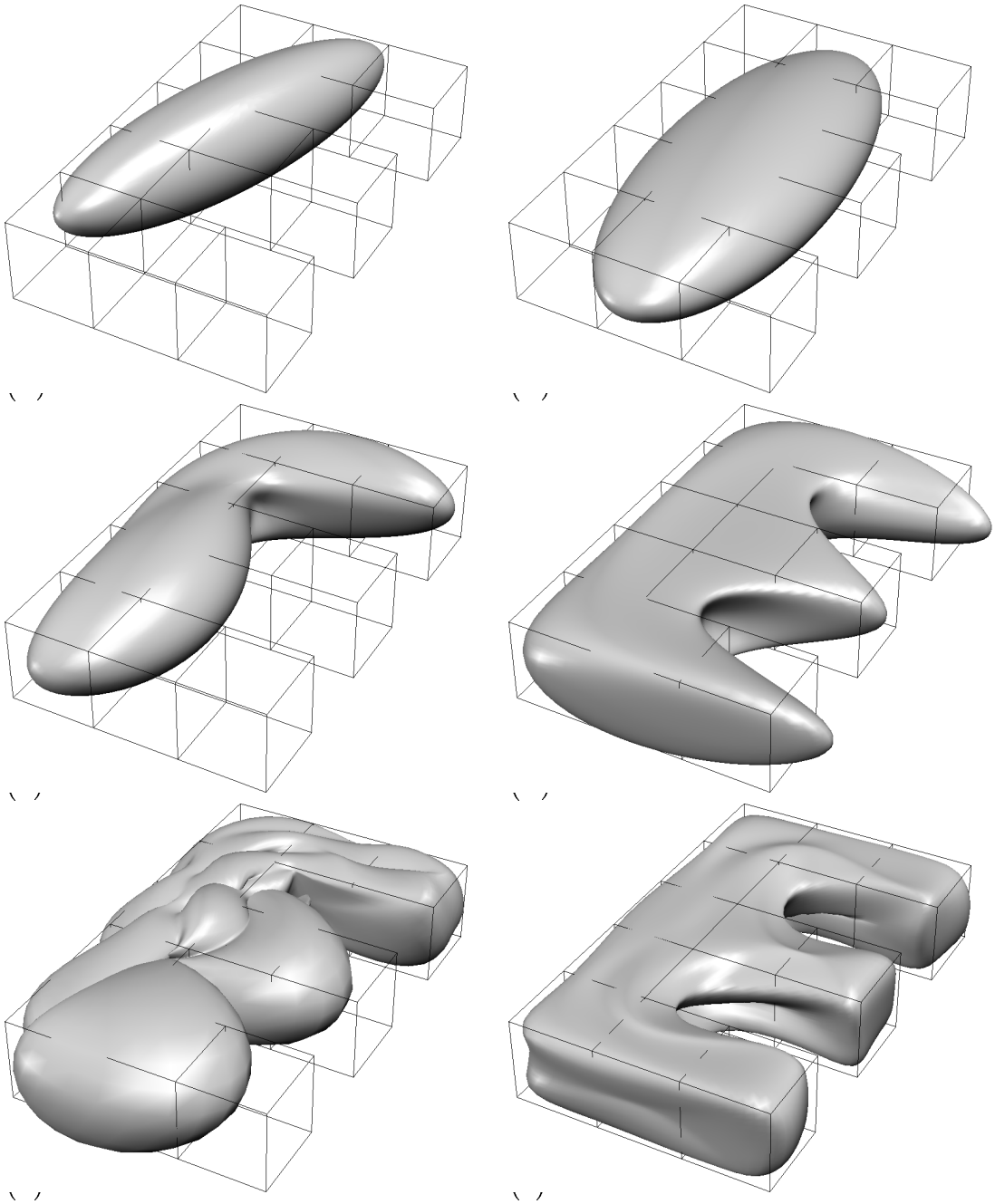


Figure 5.19: Homogeneous parameter distribution is important for shape description. The "E"-shaped object surface, indicated by a wireframe, is expanded into a series of spherical harmonics. The Shaded surfaces depict the reconstructions of the series up to degree 1 (a,b), 4 (c,d), and 10 (e,f). The initial, non-uniform parametrization yields a poor shape representation (a,c,e); its optimization achieves a significant improvement (b,d,f). $\|E\|_F$ measures the error quantitatively; $n_{\text{vert}}^{-1/2}\,\|E\|_F$ gives the RMS distance in pixel units (a: 1.143, b: 0.884; c: 0.729, d: 0.250; e: 0.313, and f: 0.102).

5.3.1 Symmetries of the Ellipsoid

The symmetry of the first degree ellipsoid allows for four different standard orientations of the object. Each of them results from the other three by mirroring (rotating by π) about one of the three main axes. The four pose transformations form a commutative group with respect to concatenation. The group is known as Klein's four-group. The concatenation of two mirrorings results in identity if they are about the same axes, and in a mirroring about the third axis if they are different. An object can flip in this way in parameter space and in object space independently, which gives rise to sixteen combinations. In the comparison of two descriptors, the minimal distance resulting from any of the sixteen relative flips is relevant as the dissimilarity measure.

Any π flip leaves the magnitudes of all coefficients constant, however it changes the sign of some of the $c_{l\,k}^{m}$, where $0 \leq k < 3$ enumerates the object space coordinates. A mirroring (π flip) at the $x_{k'}$ axis in object space flips the sign of $c_{l\,k}^{m}$ if $k' \neq k$. For example, a flip around the x_0 axis changes the signs of all $c_{l\,1}^{m}$ and $c_{l\,2}^{m}$. The situation is less simple for flips around the parameter axes. The definition of the spherical harmonics Y_l^m (eqn. 5.3) is helpful to determine which coefficients change sign in response to parameter space flips. $({u_2}^2 - 1)^l$ is an even symmetric polynomial of degree $2l$ in u_2. Its $l+m$-th derivative is a polynomial of degree $2l - (l+m) = l - m$. This derivative is even symmetric when $l - m$ is even, and odd otherwise. Multiplying it with the even function $(1 - {u_2}^2)^{m/2}$ does not change its symmetry and yields a multiple of P_l^m. Constant factors are insignificant for the present symmetry considerations. Y_l^m (up to a constant factor) results from the multiplication by $e^{im\phi} = (u_0 + iu_1)^m$. The real part of the latter is m-symmetric in u_0 and even in u_1. The imaginary part is $(m-1)$-symmetric in u_0 and odd in u_1; it vanishes for $m = 0$. In terms of symmetry, Y_l^m is equivalent to $u_2^{l-m}\, u_0^m\, (1 + i\, u_0\, u_1)$ (only the parity of the exponents is relevant). Table 5.2 summarizes the sign changes.

The formulas in this table cause the sign change pattern that table 5.3 makes explicit. Any flip around one axis is the concatenation of the flips around the other two axes, and the condition for any one sign change is the exclusive disjunction of the other two.

Object recognition tries to match a given object with a similar one out of a collection of known model objects. If it is to use some form of descriptors, it has to describe all model objects and collect their descriptors ahead of time. The recognition then takes two steps. First it

A flip around axis	changes signs of	multiply $\mathrm{Re}(c_l^m)$ by	multiply $\mathrm{Im}(c_l^m)$ by
u_0	$u_1,\ u_2$	$(-1)^{l+m}$	$(-1)^{l+m+1}$
u_1	$u_0,\ u_2$	$(-1)^{l}$	$(-1)^{l+1}$
u_2	$u_0,\ u_1$	$(-1)^{m}$	$(-1)^{m}$

Table 5.2: The sign changes that parameter space flips cause.

l	$m=0$	1		2		3		4	
		Re	Im	Re	Im	Re	Im	Re	Im
0	–								
1	01	12	02						
2	–	02	12	–	01				
3	01	12	02	01	–	12	02		
4	–	02	12	–	01	02	12	–	01

Table 5.3: The formulas in table 5.2 evaluate to this sign change pattern. A "0", "1" or "2" entry means this coefficient changes sign when the object is mirrored on the u_0, u_1 or u_2 axis, respectively. The sign of an coefficient marked "–" will never change due to a π rotation about any parameter axis.

describes the unknown object. Then it compares that descriptor with all the descriptors of the models. If any distance is less than a preset dissimilarity threshold – or tolerance –, the unknown object is considered to match the corresponding model. A particular strategy might pick the closest match if several existed. For a large model database, indexing can give significant savings compared to the comparison with *all* model descriptors.

When the task is not only to recognize an object but also to estimate its pose, the parameters of the transformation to standard position, orientation and size must be stored along with each descriptor. The transformation has 7 degrees of freedom, divided in 3 for translation (in $\underline{c}_0^0$), 3 for rotation, and one for scale.

5.3.2 Limitations

A word of caution is in order. The descriptors introduced above are prone to serious quantization artefacts. These can lead to the undesirable situation where two copies of the same object produce different descriptors, although they are only rotated, translated and/or scaled with respect to each other. When the object is too small, i.e. of the order of a few voxel units, a sampled voxel version cannot represent it adequately. Sampling the object thus misses shape aspects that cannot be recovered from the voxel collection. This leads one kind of quantization artefacts which can be corrected by sampling the object at sufficient resolution.

The parametrization by its construction assigns exactly the same amount of parameter space to each surface facet. But there are situations that do not justify this. The same planar area on the surface of an ideal object corresponds to a varying number of facets, depending on its orientation. When the planar surface is parallel to one of the coordinate planes and is sufficiently large, it is discretized into a number of facets approximately equal to its true area, measured in square grid units. When we rotate the surface by $\pi/4$ around one of the two coordinate directions it is parallel to, its quantization turns into a stair, and the number of facets for the same area multiplies by $\sqrt{2}$. When the surface is orthogonal to $(1,1,1)^T$ (diagonal in space), the number of facets is even $\sqrt{3}$ times the true area. The various regions of an object's surface can claim different amounts of parameter space with respect to each other, depending on the rotational position of the object. This leads to a different parametrization and hence to different descriptors, which cannot made invariant. If there is an object that presents this problem, the problem will persist, no matter how fine the discretization is.

5.3.3 Experimental results

In a student project, Matteo Frapolli compared the spherical shape descriptors for a number of small test objects. Table 5.4 presents the objects he prepared for these experiments. He calculated invariant descriptors for all the test objects. Table 5.5 lists all the distances between the descriptors of any two shapes. Spherical harmonics up to degree 7 have been used, i.e. $n_l = 8$. As a metric, this matrix must be symmetric. The most similar objects are "c" and "c2", as expected. Discretization artefacts cause the remaining differences; the small objects are at the limit of the resolution.

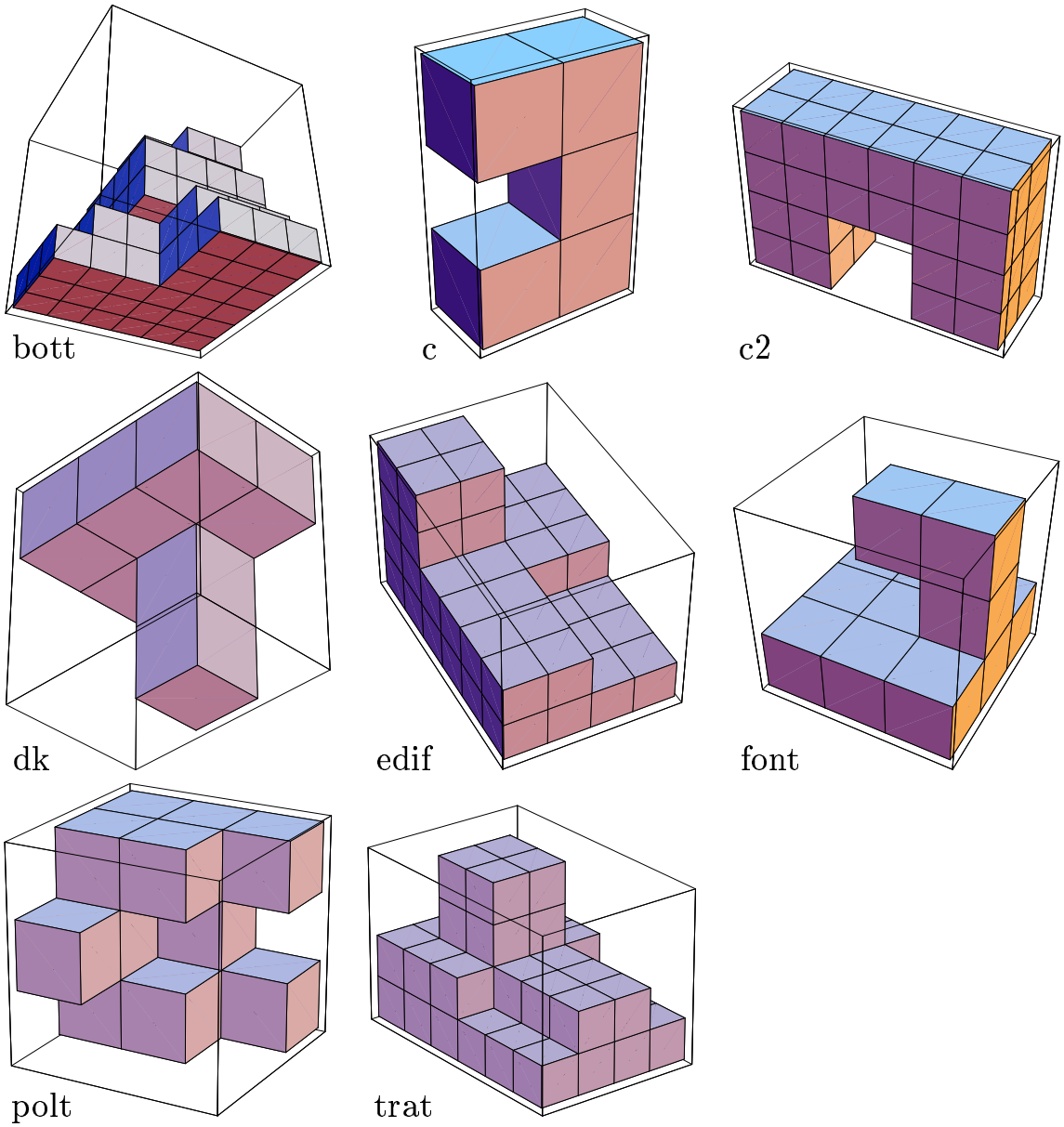


Table 5.4: A collection of simple test objects for shape comparison, together with their working names. Object "c2" is a copy of "c", enlarged by voxel replication.

	bott	c	c2	dk	edif	font	polt	trat
bott	0	0.3957	0.3910	0.1553	0.2064	0.2588	0.1759	0.2079
c	0.3957	0	0.0159	0.5251	0.1811	0.5255	0.5381	0.2388
c2	0.3910	0.0159	0	0.5296	0.1824	0.5386	0.5094	0.2412
dk	0.1553	0.5251	0.5291	0	0.3133	0.2071	0.2779	0.2884
edif	0.2064	0.1811	0.1824	0.3133	0	0.3237	0.2830	0.1132
font	0.2588	0.5255	0.5386	0.2071	0.3237	0	0.3364	0.3267
polt	0.1759	0.5381	0.5094	0.2779	0.2830	0.3364	0	0.3239
trat	0.2079	0.2388	0.2412	0.2884	0.1132	0.3267	0.3239	0

Table 5.5: Dissimilarity between the descriptors and hence between the shapes of any two of the small test objects.

Larger Objects

The surfaces of the following test objects are parametrized, and their shapes are expressed with spherical harmonic descriptors.

- *c4* A "c"-shaped polyhedron made up from five 4x4x4 voxel cubes.
- *c8* The same object, magnified by a factor of two in all coordinate directions.
- *box A* The voxels fill a rectangular box, which is not aligned with the coordinate axes.
- *box B* Rotating the box A results in a completely different sampling.
- *patella* The patella was extracted from a segmented CT scan data set of a human knee.
- *ventricle* The ventricular system has been segmented from a MRI data set of a hydrocephalus patient[3]. We selected one lateral ventricle. The data has been interpolated to compensate for the aspect ratio of 1:1:6.4 of the original data.

For each object, Figures 5.20, 5.21, and 5.22 present the cuberille interpretation of the input data, a spread-out graph of the parametrization, and the reconstruction from spherical harmonic descriptors. The cylindrical projection we chose for drawing the parametrization shows the

[3]Data courtesy Ron Kikinis, M. D., Surgical Planning Lab, Department of Radiology, Brigham and Women's hospital and Harvard Medical School, Boston

true area ratios. The smooth surface of the reconstruction is shaded in a pattern that allows the estimation of the parameter values. These latter parameter values do not coincide with the ones in the middle diagram. They rather differ by the rotation in parameter space that makes the descriptors rotation invariant. The object space rotation is suppressed in the diagrams to show the spatial relation of the original data – shown as a wireframe – with the reconstruction from the descriptor, up to degree 8. In the case of the ventricle, this shows an insufficient degree of detail, but the same value was chosen for comparability.

Table 5.6 summarizes the sizes and differences of the various test objects. Virtually all of the processing time for an object is spent in the optimization. The figure for computation time must be interpreted with caution; it qualifies only the optimization program, which is not necessarily as efficient as possible, and which might be substituted with an out-of-the box optimizer. Times are measured on a HP 9000/735. The number of vertices, n_{vert}, indicates the size of the problem: the optimization has $3\,n_{\text{vert}}$ variables, $2n_{\text{vert}} - 3$ equality constraints and $4\,n_{\text{vert}} - 8$ inequality constraints.

The distance between the descriptors appears to be a valid rough measure of shape dissimilarity. The matrix of distances is symmetric by definition. The two "c"s are most similar to each other. The two boxes are also quite similar. Both these examples illustrate the translation, rotation and scale invariance of the descriptors. The patella is more similar to a box than to a "c", whereas the ventricle is more similar to a "c" than to any of the other objects.

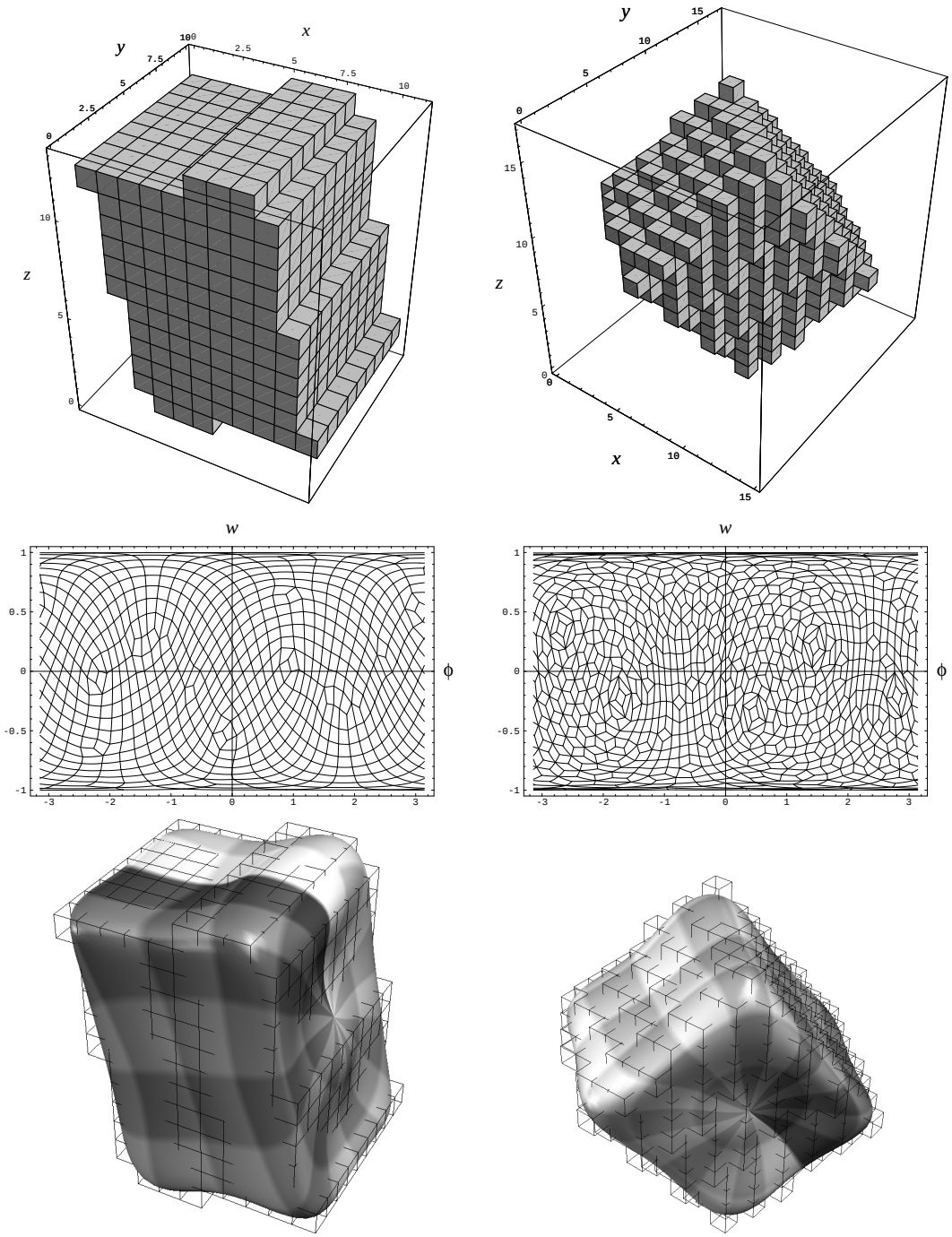


Figure 5.20: Boxes as test objects. Left column: object "box A"; right column: "box B". From top to bottom in each column: The cuberille interpretation of the input data, the parametrization drawn as a flat net spread out in $(\phi, w = \cos\theta)$ parameter space, and the reconstruction overlayed with a wireframe of the original data.

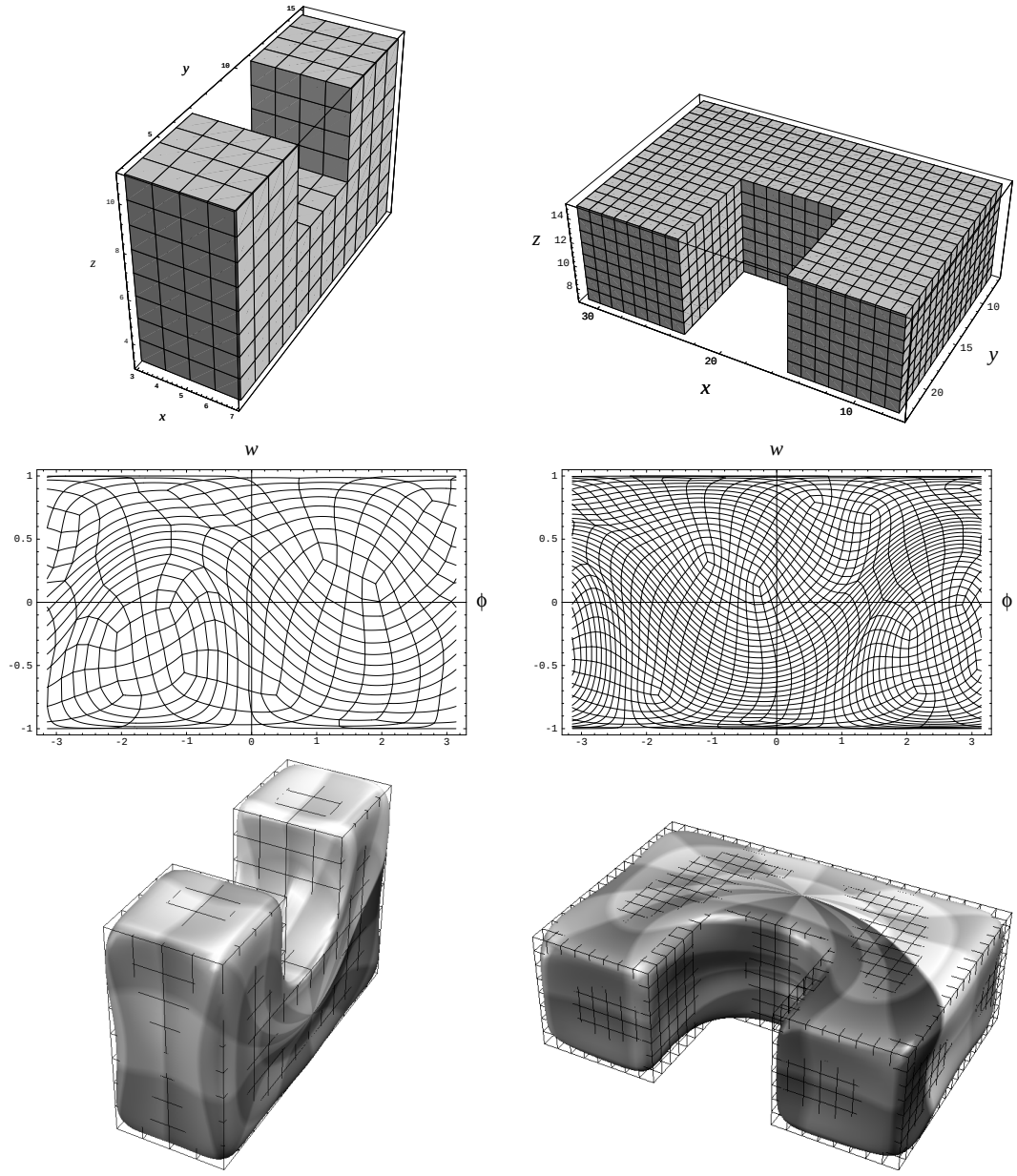


Figure 5.21: "C" letters as test objects. Left column: object "c4"; right column: "c8". From top to bottom in each column: The cuberille interpretation of the input data, the parametrization drawn as a flat net spread out in $(\phi, w = \cos\theta)$ parameter space, and the reconstruction overlayed with a wireframe of the original data.

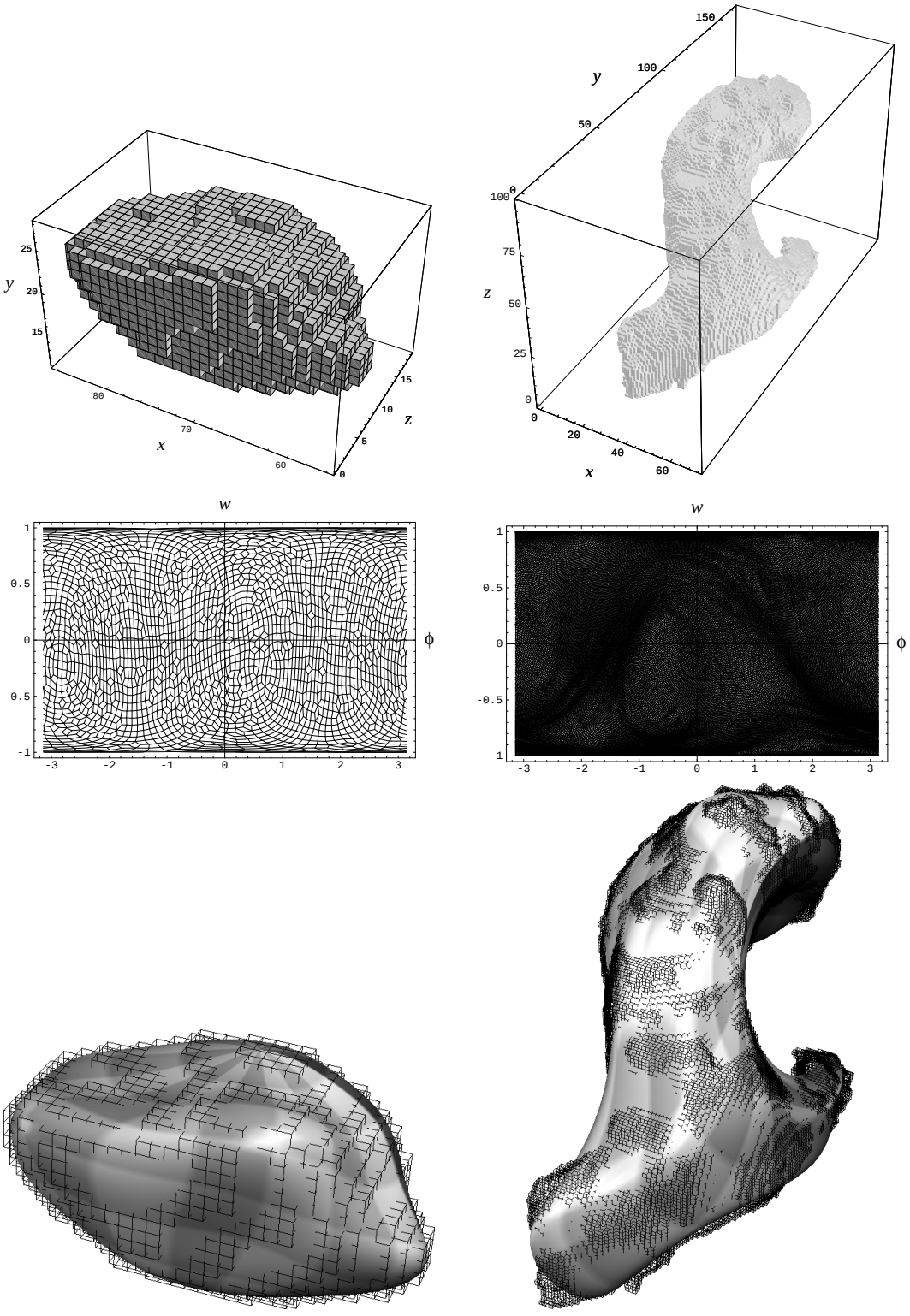


Figure 5.22: Medical test objects.Left column: "patella" object; right column: "ventricle". From top to bottom in each column: The cuberille interpretation of the input data, the parametrization drawn as a flat net spread out in $(\phi, w = \cos\theta)$ parameter space, and the reconstruction overlayed with a wireframe of the original data.

	name					
	box A	box B	c4	c8	patella	ventricle
n_{vert}	628	902	354	1410	2182	37654
time	33 s	267 s	26 s	338 s	536 s	28 h
distance to						
box A	0	0.0241	0.2370	0.2378	0.0673	0.3850
box B	0.0241	0	0.2796	0.2808	0.0859	0.4555
c4	0.2370	0.2796	0	0.0002	0.2309	0.2175
c8	0.2378	0.2808	0.0002	0	0.2299	0.2143
patella	0.0673	0.0859	0.2309	0.2299	0	0.2623
ventricle	0.3850	0.4555	0.2175	0.2143	0.2623	0

Table 5.6: Comparison of the six test objects, including the squared Euclidean distances between their descriptors.