Shape Space from Deformation *

Ho-Lun Cheng†, Herbert Edelsbrunner‡ and Ping Fu§

Abstract

The construction of shape spaces is studied from a mathematical and a computational viewpoint. A program is outlined reducing the problem to four tasks: the representation of geometry, the canonical deformation of geometry, the measuring of distance in shape space, and the selection of base shapes. The technical part of this paper focuses on the second task: the specification of a deformation mixing two or more shapes in continuously changing proportions.

1 Introduction

Geometric shapes populate our 3-dimensional physical world in a seemingly inexhaustible variety. In his famous treatise, Riemann characterizes the space of all shapes as an infinite-dimensional manifold [20]. The variety precipitates in entire mathematical disciplines focussed on subclasses of shapes, such as convex bodies in convex geometry [10], smooth manifolds in differential geometry [11], self-similar shapes in fractal geometry [17], etc. This paper takes initial steps towards an algorithmic treatment of geometric shapes and the space they define. By introducing a canonical deformation between shapes, we define and construct low-dimensional spaces of shapes. These can be viewed as subspaces of Riemann's infinite-dimensional shape manifold. The eventual goal is a computer system that supports a broad range of shape manipulation mechanisms, including creation, deformation, approximation, search, animation, and analysis. To motivate the particular approach taken in this paper, we consider work and problems in three related areas: biological shape variation, geometric morphing, and structural molecular modeling.

*This research is partially supported by the National Science Foundation under grants CCR-96-19542 and CCR-97-12088, and by the Army Research Office under grant DAAG55-98-1-0177.

†Department of Computer Science, University of Illinois, Urbana, Illinois.

‡Department of Computer Science, Duke University, Durham, and Raindrop Geomagic, Research Triangle Park, North Carolina.

§Raindrop Geomagic, Research Triangle Park, North Carolina.

Biological shape variation. Morphometrics is a quantitative study of biological shape and its variation. The theory is based on landmark points marking important features; see Small [23] but also Brookstein [3]. The sequence of landmark points defines an index into a possibly high-dimensional manifold where shapes are points and their distance is measured by the Procrustean metric. The reliance on subjectively identified points is of course problematic. Another and possibly more severe limitation of the landmark approach results from the high dimension needed to capture a reasonable amount of detail of even rather simple shapes.

In the approach taken in this paper we construct a space from a collection of base shapes. The dimension of the space depends only on the number of base shapes and not on the amount of detail or complication they represent. For many natural classes, such as for example the class of human faces, it should be possible to index a shape with a fairly small number of coordinates. How small a number suffices and how well the indexing scheme works depends on the richness of the class and on our ability to identify the types that span all or most members in the class.

Geometric morphing. In computer graphics the gradual change of a source shape into a target shape is referred to as *metamorphosis* [12] or *morphing* [13]. The primary objective in this field is the generation of pictures. A pragmatic consequence is that images are more important than geometry: the computation of a shape is avoided if convincing pictures can be generated without it. Image morphing is considerably easier than shape morphing, and the last few years have witnessed the widespread use of image morphing techniques in movie production and advertisement. However, to produce images of shapes that change during motion it will be necessary to apply the morphing process directly to shapes. In contrast to the work in computer graphics, this paper focuses on the act of deformation and side-steps the problem of establishing correspondences used in guiding the deformation.

An important question in any attempt to geometric modeling and morphing is how shapes are represented. In this paper we settle on the skin representation introduced in [7].

This is similar to but different from the blobby description introduced more than 15 years ago by Blinn [2]. That method constructs a density function $\mathbb{R}^3 \to \mathbb{R}$ as the sum of base functions or *blobs*, and it defines a shape through a level set describing its boundary. The method has met some commercial success as the metaball technique [19]. An extension of the blobby method to morphing has recently been described in [9].

Structural molecular modeling. One of the traditional models in biochemistry represents an atom by a ball and a molecule by the geometric union of atom balls. Variations of this basic idea have been developed a quarter century ago by Lee and Richards [15]. Depending on the choice of radii, the union of balls is referred to as the *van der Waals* or the *solvent accessible* model of the molecule. The skin of the set of atom balls as described in Section 3 suggests itself as an alternative geometric model. It is the only one tangent continuous at all points, which is a property sometimes falsely claimed for the *molecular surface* model obtained by rolling a solvent sphere over the van der Waals model [5].

In the field of molecular dynamics the motion of large assemblies of molecules is simulated based on the analysis of local and global forces [16]. Dramatic improvements both in efficiency and in accuracy will be needed to produce a computational tool that reliably simulates complication motions, such as for example the folding of a protein [21]. The ability to predictably deform, following physical forces or not, is also useful in the comparative study of different conformations of the same molecule. Differences can be bridged by the automatic deformation developed in later sections of this paper. The resulting continuous evolution establishes a temporal framework in which difference and similarity can be objectively described.

Outline. Section 2 describes a programmatic approach to computing shape spaces. Section 3 reviews the skin representation of shapes, which is one of the cornerstones of our method. Sections 4, 5, 6 explain the canonical deformation between shapes, which is another cornerstone. Section 7 concludes this paper by sketching further steps towards a computational solution to the shape space problem.

2 The Program

A method for the construction of concrete and low-dimensional shape spaces is outlined. We begin with the general framework and reduce the problem to four major computational tasks.

General framework. For the purpose of explaining the general principle of shape space creation, we define a *shape* as a subset of some finite-dimensional Euclidean space. For specificity but without implied restriction of generality let this space be $\mathbb{R}^3$. Given two *base shapes*, $\mathbb{X}_0, \mathbb{X}_1 \subseteq \mathbb{R}^3$, we create a 1-dimensional segment of shapes which we write as

$$\mathbb{X}_t = (1-t) \cdot \mathbb{X}_0 + t \cdot \mathbb{X}_1, \tag{1}$$

for $t \in [0, 1]$. If we have $m + 1$ shapes, $\mathbb{X}_0, \mathbb{X}_1, \ldots, \mathbb{X}_m \subseteq \mathbb{R}^3$, we define an m-dimensional simplex of shapes:

$$\mathbb{X}_T = \sum_{i=0}^{m} t_i \cdot \mathbb{X}_i, \tag{2}$$

where $T = (t_0, t_1, \ldots, t_m) \in \mathbb{R}^{m+1}$ with $\sum_{i=0}^{m} t_i = 1$ and $t_i \geq 0$ for every i. The *time vector*, T, generalizes the standard 1-dimensional notion of time. Imagine $\mathbb{X}_T$ as a point inside the infinite-dimensional manifold of all shapes in $\mathbb{R}^3$, see Figure 1. While the computational representation

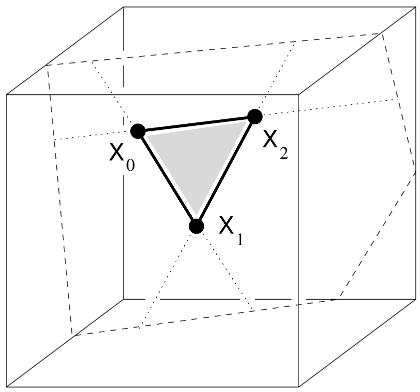

Figure 1: The cube symbolizes the infinite-dimensional manifold of all shapes. Three base shapes define a 2-dimensional triangle of shapes. Every point in the triangle is a shape obtained by combining the base shapes in unique proportions.

of the infinite-dimensional manifold of shapes seems hopelessly out of reach, the construction of low-dimensional subspaces is feasible, as demonstrated in this paper.

Task reduction. To make the above abstract approach concrete, we formulate four tasks that amount to a computational solution to low-dimensional shape spaces. The key term here is 'concrete' with the eventual goal being a working computer system for shape manipulation.

I A uniform **representation** of shapes forms the foundations of the system.

II A canonical **deformation** of shapes gives meaning to formulas (1) and (2).

III A **metric** aids in the approximation of shapes outside by shapes inside the constructed space.

IV A collection of **base shapes** spans a space that contains or approximates every shape in a class.

We build our system on the shape representation using sphere and blending patches as described in [7]. The essential features of that representation are explained in Section 3. A canonical deformation between two or more shapes is developed in this paper. Section 4 discusses the relatively easy deformations implied by simultaneous local growth. Section 5 extends the ideas to include deformations implied by simultaneous local motion. Section 6 extends the construction from 2 to $m + 1 > 2$ shapes. Section 7 sketches approaches to Tasks III and IV.

3 Representing Geometry

This section reviews the shape representation described in [7]. It consists of a simplicial complex capturing structure and connectivity and a smooth surface used for form and appearance.

An example. Figure 2 shows the complex of a shape defined by five spheres. It consists of five vertices, six edges, and one triangle. Observe the tunnel passing through the frame of four edges. Figure 3 shows a corresponding smooth surface referred to as *skin*. The five spheres are still visible and connected via blending patches of hyperboloids. Note that the complex and the body bounded by the skin are connected the same way, with a single tunnel through the middle of the shape. While the complex is a combinatorial structure, the skin is the smooth surface of a geometric shape.

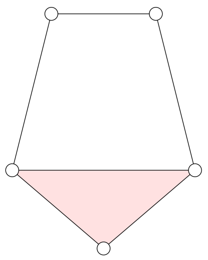

Figure 2: The complex of a shape defined by five spheres in 3-dimensional space.

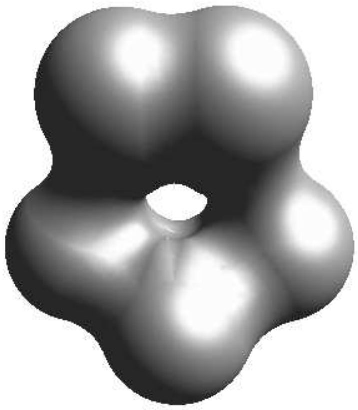

Figure 3: The skin that corresponds to the complex in Figure 2.

Geometric properties. The similarity between the complex and the skin in Figures 2 and 3 is not coincidental. Some of the properties of complex and skin that explain similarities and differences are stated below. Proofs are omitted and can be found in [7].

(P1) In the generic case, the skin is tangent continuous.

This means that for every point x on the skin surface there is a unique tangent plane passing through x. The tangent plane passing through another point y of the skin approaches the plane of x if y approaches x.

(P2) The skin consists of finitely many patches, each part of a quadric surface.

More specifically, each patch is either a piece of a sphere or a piece of a hyperboloid of revolution. The patches meet in circular and hyperbolic arcs that lie in planes separating adjacent patches.

(P3) The body bounded by the skin has the same homotopy type as the complex.

In plain English this means the body and the complex are connected the same way: they have the same number and arrangement of components, tunnels, and voids.

(P4) The skin is symmetric with respect to inside and outside.

In other words, exactly the same skin surface can be defined from two sides, by a finite set of balls inside the surface and another set outside the surface. The outside set is uniquely defined by the inside set and vice versa.

Definition of complex. Let $B = \{b_1, b_2, \ldots, b_n\}$ be a set of closed balls in $\mathbb{R}^3$. We write $b_i = (z_i, \varrho_i)$, where z_i is the center and ϱ_i is the radius of b_i. The union, $\bigcup B$, is the set of points $x \in \mathbb{R}^3$ contained in at least one of the balls, with and example shown in Figure 4. It is generally a non-convex set bounded by sphere patches that meet along circular arcs. The arcs meet at corner points where three or more patches come together. The *weighted distance* of a point x from a ball b_i is

$$\pi_i(x) = \|x - z_i\|^2 - \varrho_i^2,$$

which is positive outside, zero on the boundary, and negative in the interior of b_i. We decompose $\bigcup B$ into convex sets using (weighted) Voronoi cells. Specifically, the *Voronoi cell* of

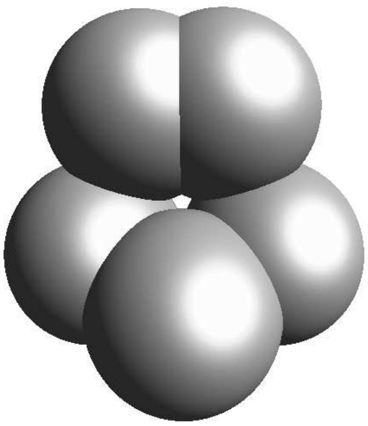

Figure 4: The five spheres that define the complex in Figure 2 and the skin in Figure 3.

b_i is the set of points x for which b_i minimizes the weighted distance:

$$V_i = \{x \in \mathbb{R}^3 \mid \pi_i(x) \le \pi_j(x), 1 \le j \le n\}.$$

The Voronoi cells are convex polyhedra that cover the entire $\mathbb{R}^3$. The intersection of a Voronoi cell with the union of balls is $V_i \cap \bigcup B = V_i \cap b_i$, which is convex as anticipated.

The complex defined by B is dual to the complex of cells $V_i \cap b_i$. For generic sets B, it consists of simplices only, but genericity is not a feasible assumption in a context where the deformation is based on persistent non-generic positions of balls. In this more general setting, the complex consists of convex polytopes of dimension 0 through 3. A convex polytope of dimension k is referred to as a *k-polytope*. For example, 0-polytopes are vertices and 1-polytopes are edges. Each k-polytope is the convex hull of $\ell+1 \ge k+1$ ball centers. Specifically, the convex hull of centers z with indices $i_0, i_1, \ldots, i_\ell$ is a k-polytope in the complex iff the common intersection of the corresponding $\ell+1$ balls and Voronoi cells, $\bigcap_{j=0}^{\ell}(V_{i_j} \cap b_{i_j})$, has dimension $3-k$ and no other Voronoi cell contains this set. We denote the complex as $\operatorname{Dsx} B$ to indicate it is a subcomplex of the Delaunay complex and it is defined by B; see also [6, 8].

Definition of skin. The skin is defined as the envelope of an infinite family of spheres. The family is generated by adding, scaling, and shrinking spheres in a given finite set. Think of a sphere as the zero-set of the weighted distance function and a ball as the union of concentric spheres: $b_i = \pi_i^{-1}(-\infty, 0]$. The sum of two functions and the multiplication with a scalar are defined as usual:

$$\begin{aligned}(\pi_i + \pi_j)(x) &= \pi_i(x) + \pi_j(x),\\ (\gamma \cdot \pi_i)(x) &= \gamma \cdot \pi_i(x).\end{aligned}$$

The collection of weighted distance functions together with addition and scaling forms a vector space. Let $\Pi = \{\pi_i \mid b_i \in B\}$ be the set of weighted distance functions of the given balls. The affine hull and the convex hull of Π are defined as usual:

$$\begin{aligned}\operatorname{aff}\Pi &= \{\pi = \sum_{i=1}^{n} \gamma_i \cdot \pi_i \mid \sum_{i=1}^{n} \gamma_i = 1\},\\ \operatorname{conv}\Pi &= \{\pi \in \operatorname{aff}\Pi \mid \gamma_i \ge 0 \text{ for all } i\}.\end{aligned}$$

The pointwise minimum of the functions in $\operatorname{conv}\Pi$ is again a function $\mathbb{R}^3 \to \mathbb{R}$. Its zero-set is the boundary of $\bigcup B$. The final step of the construction shrinks every sphere by a factor of $\sqrt{2}$ towards its center. Let $\pi(x) = \|x-z\|^2 - \varrho^2$ and define $\pi'(x) = \|x-z\|^2$. We formalize the shrinking operation by defining $\tilde{\pi}(x) = \pi(x) + \pi'(x)$. Indeed, the zero-set of $\tilde{\pi}$ is the set of points x that satisfy $\|x-z\|^2 - \varrho^2/2 = 0$. This is the sphere with center z and radius $\varrho/\sqrt{2}$. The *skin* is formally defined as the zero-set of the function obtained by taking the pointwise minimum over all shrunken weighted distance functions in the convex hull of Π, and the *body* is the part of space bounded by the skin:

$$\begin{aligned}\tilde{\Pi}(x) &= \min\{\tilde{\pi}(x) \mid \pi \in \operatorname{conv}\Pi\},\\ \operatorname{skin} B &= \tilde{\Pi}^{-1}(0),\\ \operatorname{body} B &= \tilde{\Pi}^{-1}(-\infty, 0].\end{aligned}$$

Equivalently, the skin is the envelope of the infinitely many spheres that are the zero-sets of the $\tilde{\pi}$, $\pi \in \operatorname{conv}\Pi$. Figure 3 shows the skin defined by the five spheres in Figure 4. Figure 6 shows skin surfaces defined by one, two, three, and four spheres.

4 Growth

The simultaneous growth of all balls in B implies a restricted form of deformation for the skin. In spite of the limitation, the deformation exhibits many of the characteristics observed in the general case. The topology changes follow the case analysis common in Morse theory [18, 24].

Growth model. We choose a growth model that leaves the Voronoi cells unchanged. Most other things being equal, it has the advantage over other models that the cost for its simulation is negligible. For every $\alpha^2 \in \mathbb{R}$ let

$$b_i(\alpha) = (z_i, \sqrt{\varrho_i^2 + \alpha^2})$$

be a ball concentric to b_i. For $\alpha = 0$ the radius of $b_i(\alpha)$ is ϱ_i, and for $\varrho_i = 0$ the radius is α. For a given $\alpha^2 \in \mathbb{R}$, the set of balls is $B(\alpha) = \{b_i(\alpha) \mid b_i \in B\}$, the complex is $\operatorname{Dsx} B(\alpha)$, and the skin is $\operatorname{skin} B(\alpha)$.

We admit negative values of α^2, which correspond to imaginary α, and even negative values of $\varrho_i^2 + \alpha^2$ are allowed. In the latter case, $b_i(\alpha)$ has imaginary radius and is referred to as an *imaginary* ball. Such balls are an integral part of our theory of deformation and lend structure to the complement space. This is less apparent in the limited form of deformation implied by local growth than in the more general case implied by local motion discussed in the next section.

Change in complex topology. As α^2 increases, the cells $V_i \cap b_i$ either grow or stay the same. It follows that the complex can only gain polytopes but not lose any. Let $\alpha_1^2 \leq \alpha_2^2$, $K_i = \mathrm{Dsx}\, B(\alpha_i)$ for $i = 1, 2$, and observe that $K_1 \subseteq K_2$. Suppose K_2 contains a single polytope, σ, that does not also belong to K_1, and let σ be the convex hull of centers z with indices $i_0, i_1, \ldots, i_\ell$. There are four cases depending on the dimension k of σ. In each case the appearance of σ corresponds to the balls with indices i_0 through i_ℓ developing a common overlap for the first time. For $k = 0$ this means that b_{i_0} passes from imaginary to real radius. The four generic cases where σ is a simplex of dimension k are illustrated in Figure 5. Define a *void* as a component of the part of $\mathbb{R}^3$ not covered by the complex or the body. This includes the unbounded component, which is also called a void.

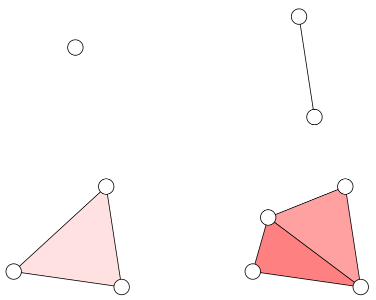

Figure 5: A new vertex, edge, triangle, tetrahedron appears as part of the complex.

CASE $k = 0$. σ is a vertex that forms a new component by itself.

CASE $k = 1$. σ is an edge that either connects two components or two portions of one component.

CASE $k = 2$. σ is a 2-polytope that either splits a void or closes a tunnel between two portions of the same void.

CASE $k = 3$. σ is a 3-polytope that fills a void.

Change in skin topology. By property (P3), the topology of the skin changes at the same time as that of the complex. Furthermore, the topology of the body bounded by the skin changes the same way as that of the complex. In other words, for each of the above four cases there is a corresponding case that describes the change in skin topology. The cases are illustrated in Figure 6. The case analysis makes reference to the local surface orientation of a patch. By this we mean the sense that distinguishes inside from outside. The body bounded by the skin consists of all points inside the skin.

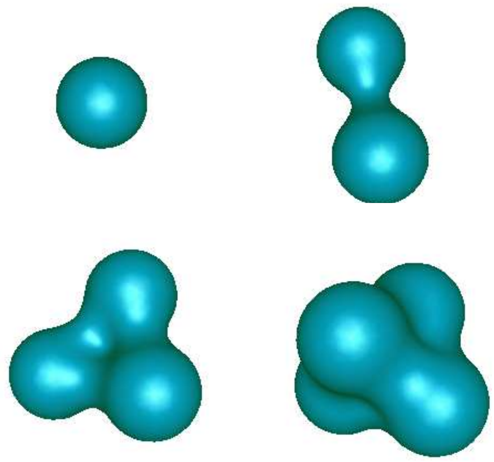

Figure 6: Cases $k = 0, 1$ at the top and $k = 2, 3$ at the bottom.

CASE $k = 0$. A component is born. The component starts out as a point that grows into a ball and eventually assumes more complicated shapes.

CASE $k = 1$. A bridge is completed. Geometrically, a hyperboloid changes from two sheets to one sheet. The two sheets approach the limiting double cone and then flip over to form one sheet. The bridge either connects two components of the skin or it connects two portions of one component.

CASE $k = 2$. A tunnel is closed. This case is symmetric to $k = 1$: a hyperboloid with opposite local surface orientation changes from one sheet to two sheets. The closing tunnel either splits a void into two or it removes a tunnel from a multiply connected void.

CASE $k = 3$. A void is filled. The void is a component of the space outside the skin surface that disappears due to the expansion of the skin. The case is symmetric to $k = 0$, with time and local surface orientation reversed.

The above case analysis mentioned the topological symmetry between the Cases $k = 0$ and 3 and between the Cases $k = 1$ and 2. Because of Property (P4), that symmetry can be observed even in the geometric detail how the topology changes happen.

5 Motion

This section focuses on a one-parametric deformation between two shapes, of which the growth model of Section 4 is a special case. In spite of the greater generality of the motion, the types of topology changes are the same as before.

Matching and interpolation. Let B and C be two finite sets of balls in $\mathbb{R}^3$. The two sets define two shapes and we are interested in their body representations: $\mathbb{X}_0 = \mathrm{body}\, B$

and $\mathbb{X}_1 = \text{body } C$. Intermediate shapes are construction by interpolation between B and C. It is convenient to project a cross-section of the vector space of weighted distance functions onto the set of balls. Formally, if b and c are balls with weighted distance functions π and φ and $\gamma_1 + \gamma_2 = 1$ then $a = \gamma_1 \cdot b + \gamma_2 \cdot c$ is the ball with weighted distance function $\gamma_1 \cdot \pi + \gamma_2 \cdot \varphi$. With this introduction define

$$\begin{aligned} A_t &= (1-t) \cdot B + t \cdot C \\ &= \{(1-t) \cdot b + t \cdot c \mid b \in B, c \in C\}, \end{aligned}$$

and consider the one-parametric family of bodies $\mathbb{X}_t = \text{body } A_t$ defined for all $t \in [0, 1]$. Figure 8 at the end of this paper illustrates the definition by showing the skin surface of a hexagonal ring at $t = 0.0$ deforming to a half-circle with bottom at time $t = 1.0$. Observe that the construction of A_t is independent of location and orientation in space. In other words, if $\tau : \mathbb{R}^3 \to \mathbb{R}^3$ is a rigid motion then $\tau(A_t) = (1-t) \cdot \tau(B) + t \cdot \tau(C)$.

At any time t in the open interval between 0 and 1, the set A_t contains a ball for every pair $(b, c) \in B \times C$. The complete matching avoids the difficulty of determining a correspondence between the various portions or features of $\mathbb{X}_0$ and $\mathbb{X}_1$; such a correspondence will automatically be established, although it is not bijective. It seems that by definition the number of balls in A_t is the product of the numbers for B and C, but the geometry of the construction causes many of these balls to be redundant, in the sense that their Voronoi cells are empty. Which of the balls $a_t = (1-t) \cdot b + t \cdot c$ are redundant depends on the relative distance between b and c. If the Voronoi cells of a_t is empty for some $t \in (0, 1)$ then it is empty for every such t, and there are algorithms that construct $\text{Dsx } A_t$ spending time only on non-redundant pairs (b, c). These algorithms are based on a reinterpretation of the construction as a convex hull in $\mathbb{R}^5$ described shortly. A particular such output-sensitive convex hull algorithm can be found in [22]. However, the structure of the special 5-dimensional problem permits a simpler algorithm that can be viewed as overlaying two 3-dimensional Voronoi complexes, which can be done in logarithmic time per cell. This algorithm has been implemented by the authors of this paper and is used in the animation of deformations.

Trading dimensionality for convexity. We identify $\mathbb{R}^3$ with the linear subspace spanned by the first three coordinates of $\mathbb{R}^5$. The fourth coordinate is used to turn 3-dimensional non-convex shapes into 4-dimensional convex shapes, and the fifth coordinate is used to cast dynamic change over time into static geometry. To turn non-convex into convex geometry, we interpret a shape in $\mathbb{R}^3$ as the projection of the intersection of two convex shapes in $\mathbb{R}^4$. More specifically, one of the two shapes is a 4-dimensional convex body and the other is a convex surface bounding a 4-dimensional convex body:

CONVEXIFICATION PRINCIPLE.

$$\text{shape}^3 = \text{proj}\,(\text{conv}^4 \cap \text{bd conv}^4).$$

The principle applies to the definition of skin; see [7]. In this case the convex surface is the graph of the distance square function from the origin: $\pi_0 : \mathbb{R}^3 \to \mathbb{R}$ defined by $\pi_0(x) = \|x\|^2$. The convex body is derived from the convex hull of the set B lifted to 4 dimensions:

$$\lambda(B) = \{(z_i, \|z_i\|^2 - \varrho_i^2) \in \mathbb{R}^4 \mid (z_i, \varrho_i) \in B\}.$$

Specifically, the body is obtained by modifying $\mathcal{B} = \text{conv } \lambda(B)$ in a way that corresponds to shrinking $\text{conv } \Pi$ as described in Section 3.

Most important for the reinterpretation of the deforming construction is that the skin of B is completely specified by a 4-dimensional convex polytope, $\mathcal{B} = \text{conv } \lambda(B)$. Similarly, $\text{skin } C$ is completely specified by a 4-dimensional convex polytope, namely $\mathcal{C} = \text{conv } \lambda(C)$. Now imagine $\mathcal{B}$ and $\mathcal{C}$ embedded in parallel affine subspaces $x_5 = 0$ and $x_5 = 1$ in $\mathbb{R}^5$. Take the convex hull of the two polytopes, which is a 5-dimensional convex polytope,

$$\begin{aligned} \Xi &= \text{conv } (\mathcal{B} \cup \mathcal{C}) \\ &= \text{conv } (\lambda(B) \cup \lambda(C)), \end{aligned}$$

as illustrated in Figure 7. The set A_t corresponds to the cross-section of Ξ at $x_5 = t$. This is a 4-dimensional convex

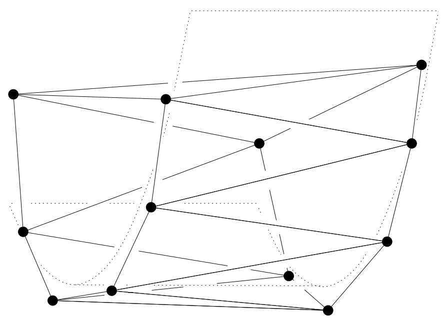

Figure 7: Sets of intervals in 1 dimension are lifted to 2-dimensional convex polygons embedded in parallel planes in $\mathbb{R}^3$. The deformation happens while a plane sweeps the convex hull of the two polygons.

polytope, namely $\mathcal{A}_t = \text{conv } \lambda(A_t)$. We apply the convexification principle to each $\mathcal{A}_t$, $t \in [0, 1]$, and thus recover the sequence of 3-dimensional shapes interpolating between the skins of B and C.

Before the beginning and after the end. The shape deformation through mixing B and C can be extended beyond $[0, 1]$ by generalizing the definition of A_t. Indeed, $A_t = (1-t) \cdot B + t \cdot C$ is well defined for all $t \in \mathbb{R}$. That this is not a very satisfying extension should be clear from Figure 7. Each pair $(b, c) \in B \times C$ is represented by a line and A_t corresponds to the cross-section of the collection of lines at $x_5 = t$. As t increases beyond 1 and goes to ∞ the lines grow apart and define progressively more spread out point sets. The same is true as t decreases below 0 and goes to $-\infty$. Correspondingly, shapes in $\mathbb{R}^3$ get bulkier and bigger. A more appropriate extension uses only balls $(1-t) \cdot b + t \cdot c$ that correspond to non-redundant combinations for values of t in $(0, 1)$. Instead of all $\operatorname{card} B \cdot \operatorname{card} C$ lines, this idea uses only lines that define edges of Ξ. An even more conservative generalization redefines Ξ as the intersection of all closed half-spaces in $\mathbb{R}^5$ that contain $\lambda(B) \cup \lambda(C)$ and whose bounding hyperplanes pass through at least one point each of $\lambda(B)$ and $\lambda(C)$ and through at least four points in total.

6 Time Vectors

This section generalizes the construction of Section 5 from two to $m + 1 \geq 2$ shapes. The resulting framework consists of a space of shapes in which deformations are interpreted as paths.

General framework. Let $B_0, B_1, \ldots, B_m$ each be a finite set of balls in $\mathbb{R}^3$. Each set defines a shape, and we consider the body representations $\mathbb{X}_i = \operatorname{body} B_i$. Let $T = (t_0, t_1, \ldots, t_m) \in \mathbb{R}^{m+1}$ with $\sum_{i=0}^{m} t_i = 1$ and $t_i \geq 0$ for each i and define

$$A_T = \sum_{i=0}^{m} t_i \cdot B_i.$$

The corresponding shape is $\mathbb{X}_T = \operatorname{body} A_T$. Note that $\mathbb{X}_T = \mathbb{X}_i$ if the only non-zero component of T is $t_i = 1$. In general, A_T contains a ball for each $(m + 1)$-tuple in $B_0 \times B_1 \times \ldots \times B_m$. In the typical case, only a small fraction of the balls in A_T are non-redundant. We take advantage of this observation and compute $\mathbb{X}_T$ without explicitly constructing the set A_T. An extension of the ideas in Section 5 maps the $m + 1$ shapes in $\mathbb{R}^3$ to a convex polytope in $\mathbb{R}^{4+m}$. The fourth coordinate realizes the convexification principle and turns 3-dimensional non-convex shapes into 4-dimensional convex polytopes. The last m coordinates represent the space of time vectors.

The above construction defines an m-dimensional simplex of shapes. If we drop the non-negativity requirement for the t_i we get an m-dimensional flat. Another meaningful extension of the shape space allows each shape to grow and shrink following a parameter $\alpha^2 \in \mathbb{R}$. As explained in Section 4, changing the value of α^2 is computationally inexpensive. With these extensions we have a shape for every $(T, \alpha^2) \in \mathbb{R}^{m+1}$. In other words, the space spanned by $m + 1$ base shapes is isomorphic to $\mathbb{R}^{m+1}$. The extra parameter, α^2, can be exploited to maintain certain properties during the deformation, such as for example the shape volume or the surface area. We refer to the companion paper [4] where the 2-dimensional space of shapes spanned by two 2-dimensional base shapes is explored in some detail.

Space and paths of shapes. The above framework associates a shape with each time vector T. In other words, T is as an index into a continuous space of shapes defined by $m + 1$ given shapes. We are more specific about this space and the parameterization through time vectors. Consider the space of all time vectors, which is isomorphic to an m-simplex. Each T in this space defines a shape $\mathbb{X}_T$, and we define $\mathbf{\Omega}$ as the space of shapes defined by time vectors. A deformation is a path $\varphi : [0, 1] \to \mathbf{\Omega}$. The simplest kind of deformation is a straight path connecting the initial with the terminal shape. Consider for example the construction in Section 5. We have $m + 1 = 2$ and $\mathbf{\Omega}$ is isomorphic to a closed line segment. The initial and terminal shapes are given by time vectors $T_0 = (0, 1)$ and $T_1 = (1, 0)$. The deformation of $\mathbb{X}_0 = \mathbb{X}_{T_0}$ into $\mathbb{X}_1 = \mathbb{X}_{T_1}$ is defined by the straight path of shapes $\varphi(t) = \mathbb{X}_{T_t}$, with $T_t = (1 - t) \cdot T_0 + t \cdot T_1 = (t, 1 - t)$, for $t \in [0, 1]$.

7 Metric and Basis

While Sections 3 through 6 provide adequate algorithms for Tasks I and II, we still lack appropriate solutions to Tasks III and IV. This section outlines what might be the most straightforward approaches to the two tasks.

Task III: a metric. Probably the best known metric of the infinite-dimensional manifold of shapes is the Hausdorff distance. Given two shapes $\mathbb{X}_0, \mathbb{X}_1 \subseteq \mathbb{R}^3$ it is the infimum over all $\varepsilon \in \mathbb{R}$ for which each point in $\mathbb{X}_i$ has a point in $\mathbb{X}_{1-i}$ at distance at most ε:

$$h(\mathbb{X}_0, \mathbb{X}_1) = \min\{\varepsilon \in \mathbb{R} \mid \mathbb{X}_i \subseteq \mathbb{X}_{1-i} + b_\varepsilon,\ i = 0, 1\},$$

where b_ε is the ball of all $x \in \mathbb{R}^3$ at distance at most ε from the origin. It is fairly straightforward to compute h in polynomial time if the shapes are given as bodies of finite sets of balls: $\mathbb{X}_0 = \operatorname{body} B_0$ and $\mathbb{X}_1 = \operatorname{body} B_1$. How fast, as a function of $n = \operatorname{card} B_0 + \operatorname{card} B_1$, can h be computed? An algorithm that rotates and translates $\mathbb{X}_1$ to minimize the Hausdorff distance can be found in [1]. An important but difficult problem is the computation of the distance between a shape $\mathbb{X} = \operatorname{body} B$ and the space spanned by $m+1$ shapes $\mathbb{X}_i = \operatorname{body} B_i$. Since $H_{\mathbb{X}} : \mathbb{R}^m \to \mathbb{R}$ defined by

$$H_{\mathbb{X}}(T) = h(\mathbb{X}, \operatorname{body} A_T)$$

seems to lack any significant structural properties other than continuity, it is not clear how to compute the infimum of $H_{\mathbb{X}}$ at all.

Task IV: base shapes. We envision a stochastic process for the identification of base shapes. Suppose $\mathbb{Y}_0, \mathbb{Y}_1, \ldots$ is a sequence of shapes in the class of interest. For an index $i \geq 0$ let $m \leq i$ and let $\mathbb{X}_0, \mathbb{X}_1, \ldots, \mathbb{X}_m$ be a collection of base shapes so each $\mathbb{Y}_j$, $0 \leq j \leq i$, is sufficiently close to some $\mathbb{X}_T$ in the defined space. If there is a time vector $T \in \mathbb{R}^{m+1}$ such that $h(\mathbb{Y}_{i+1}, \mathbb{X}_T)$ is small then $\mathbb{Y}_{i+1}$ is reasonably represented by the space and no change in the collection of base vectors is necessary. Otherwise, we may consider substituting $\mathbb{Y}_{i+1}$ for 0 or more of the base shapes. The dimension of the space increases by at most 1. There is room for plenty of refinements and improvements. Most likely it is a mistake to choose the base shapes from the class itself, although this may be most convenient at first. The all important parameter is the number of base shapes, since every increase in the dimension implies a substantial increase in complexity of all shape manipulation operations. How can we design base shapes that produce the most economical description of the space approximating a class of shapes?

References

[1] P. J. BESL AND N. D. MCKAY. A method for registration of 3-D shapes. *IEEE Trans. Pattern Analysis Mach. Intell.* **14** (1992), 239–256.

[2] J. BLINN. A generalization of algebraic surface drawing. *ACM Trans. Graphics* **2** (1980), 235–256.

[3] F. L. BROOKSTEIN. *Morphometric Tools for Landmark Data. Geometry and Biology.* Cambridge Univ. Press, England, 1991.

[4] S.-W. CHENG, H. EDELSBRUNNER, P. FU AND K.-P. LAM. Design and analysis of planar shape deformation. This journal.

[5] M. L. CONNOLLY. Analytical molecular surface calculation. *J. Appl. Cryst.* **16** (1983), 548–558.

[6] H. EDELSBRUNNER. The union of balls and its dual shape. *Discrete Comput. Geom.* **13** (1995), 415–440.

[7] H. EDELSBRUNNER. Deformable smooth surface design. *Discrete Comput. Geom.*, **21** (1999), 87–115.

[8] H. EDELSBRUNNER AND E. P. MÜCKE. Three-dimensional alpha shapes. *ACM Trans. Graphics* **13** (1994), 43–72.

[9] E. GALIN AND S. AKKOUCHE. Blob metamorphosis based on Minkowski sums. *In* "Proc. EUROGRAPHICS", **15** (1996), 143–153.

[10] P. M. GRUBER AND J. WILLS, EDS. *Handbook of Convex Geometry, volumes A and B.* North-Holland, Amsterdam, 1993.

[11] H. HOPF. *Differential Geometry in the Large.* Springer-Verlag, Berlin, Germany, 1989.

[12] A. KAUL AND J. ROSSIGNAC. Solid-interpolating deformations: construction and animation of PIPs. *In* "Proc. Eurographics, 1991", 493–505.

[13] J. R. KENT, W. E. CARLSON AND R. E. PARENT. Shape transformation for polyhedral objects. *Computer Graphics* **26** (1992), 47–54.

[14] F. LAZARUS AND A. VERROUST. Three-dimensional metamorphosis: a survey. *The Visual Computer* **14** (1998), 373–389.

[15] B. LEE AND F. M. RICHARDS. The interpretation of protein structures: estimation of static accessibility. *J. Mol. Biol.* **55** (1971), 379-400.

[16] M. LEVITT AND A. WARSHEL. Computer simulation of protein folding. *Nature* **253** (1975), 694–698.

[17] B. B. MANDELBROT. *The Fractal Geometry of Nature.* Freeman, New York, 1983.

[18] J. MILNOR. *Morse Theory.* Princeton Univ. Press, New Jersey, 1963.

[19] S. MURAKAMI AND H. ICHIHARA. On a 3D display method by metaball technique. *J. Electronics Comm.* **J70-D** (1987), 1607–1615, (in Japanese).

[20] B. RIEMANN. Über die Hypothesen, welche der Geometrie zugrunde liegen. *Abh. Königl. Ges. Wiss. Göttingen* **13** (1868).

[21] G. D. ROSE. No assembly required. *The Sciences* **36** (1996), 26–31.

[22] R. SEIDEL. Constructing higher-dimensional convex hulls in logarithmic cost per face. *In* "Proc. 18th Ann. ACM Sympos. Theory Comput., 1986", 404–413.

[23] C. G. SMALL. *The Statistical Theory of Shape.* Springer-Verlag, New York, 1996.

[24] A. WALLACE. *Differential Topology. First Steps.* Benjamin, New York, 1968.

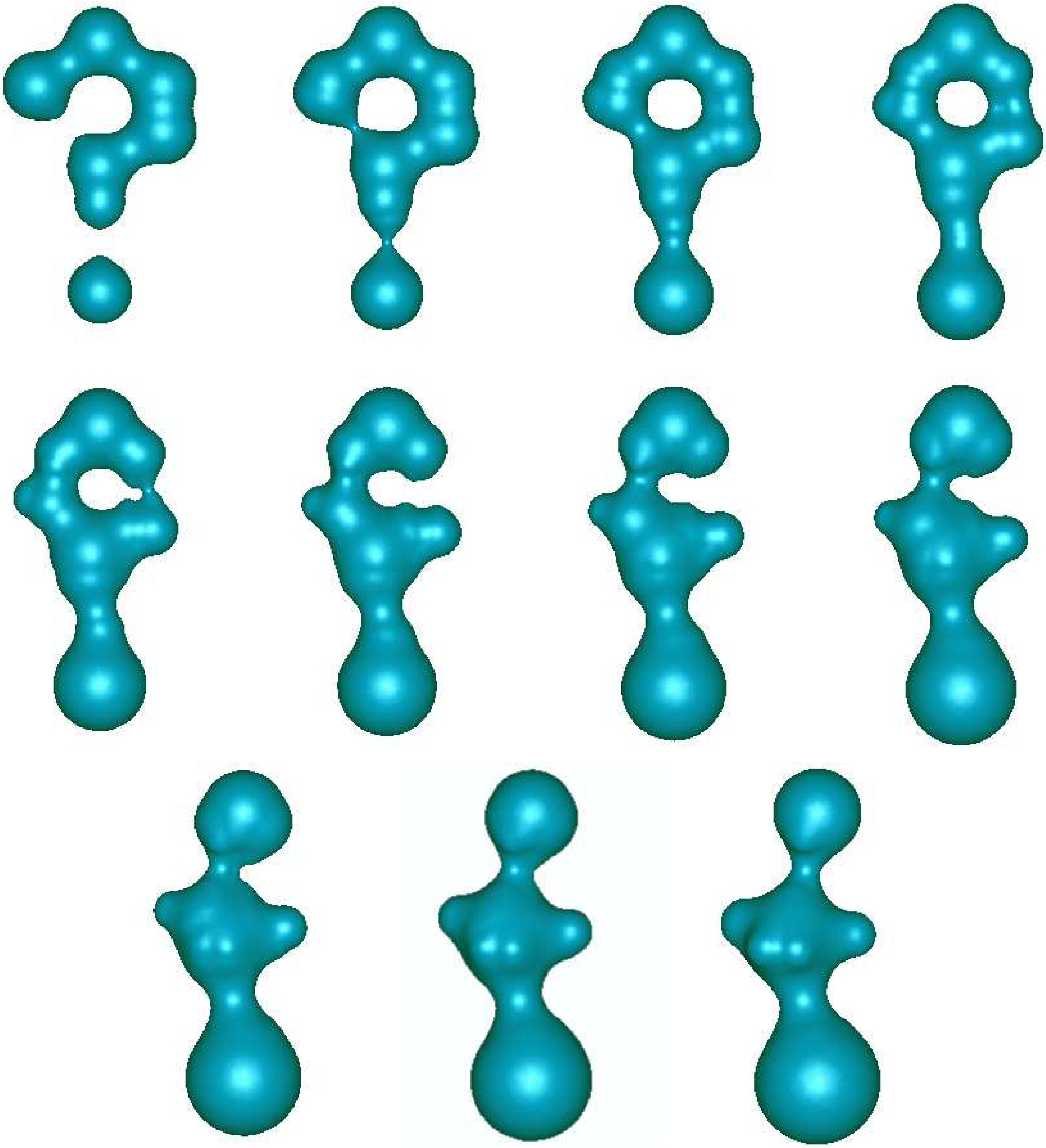

Figure 8: From left to right and top to bottom: the shapes at times $t = 0.0, 0.1, \ldots, 1.0$. The sequence is defined by a set of seven spheres forming a question mark at time $t = 0.0$ and a set of eight spheres forming a human-like figure at time $t = 1.0$.