Modal Space Subdivision for Physically-plausible 4D Shape Sequence Completion from Sparse Samples

Qing Xia¹, Shuai Li¹, Hong Qin² and Aimin Hao¹

¹State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, China
²Department of Computer Science, Stony Brook University, USA

Abstract

Subdivision techniques are powerful and ubiquitous in generating smooth surfaces from coarse 3D polygonal models, yet space-time subdivision remains unexplored when the goal is to produce dense physically-realistic shape sequence with smooth transition from a set of sparse 3D model samples. In this paper, we consider a completely data-driven strategy and propose a novel modal space subdivision scheme to facilitate the rapid generation of physically-plausible “in-between” shapes from sparse model samples. The key idea of our approach is to abstract any shape as a point in modal space, represent its potential deformation with shape modes, and iteratively seek all the intermediate shapes between two consecutive models in the modal space via recursive subdivision. Meanwhile, to guarantee deformation details, we further interpolate the shape in the local regions after global modal-space subdivision. Comprehensive experiments on various model inputs have demonstrated the power, versatility, high performance, and potential of our modal space subdivision scheme.

Categories and Subject Descriptors (according to ACM CCS): I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism—Animation

1. Introduction and Related Work

The growing wealth of static shape repositories naturally give rise to the new research topic of space-time modeling, which embraces shape deformation along the time axis. Despite limited success of many methods, however, it remains challenging to interactively make the static objects undergo physically-plausible deformations with minimum computational overhead and rapid visual feedback.

First, geometry-dominant approaches try to generate in-
termediate sequences by producing a high dimensional curve that goes through these key-models in certain geometric feature space [CH12, HSVTP12]. Even though such techniques can be implemented efficiently, the complete lack of dynamic rationales and physical correctness unavoidably induce some unpleasant artifacts, such as self-intersection and physics-contradicted unnatural shape transition.

Second, strictly following physical laws to drive one reference model to move towards the others, we could exploit physics-based methodology, whose key idea is to minimize the strain energy between adjacent key-frame models or to linearly interpolate the strain field during shape reconstruction [BSG12, HTZ*11]. However, such methods tend to be computationally demanding [BdSP09], because the underlying physical equations must be numerically integrated with implicit solvers to guarantee stability and convergence.

Third, it is also possible to consider available database with rich shape variation, and this leads to example-based approaches, whose general motivation tends to balance high physical fidelity object properties learned from examples by employing certain priors to guide the interpolation [FB11, MTGG11, GLHH13]. Although such methods aim to improve the naturalness of the intermediate shapes, the learned prior knowledge is usually model-specific and can not be directly applied to models of other categories.

![Figure 2: The shape interpolation comparison between simple linear interpolation (left) and our modal subdivision method (right) for two given woman poses.](image)

To alleviate, we recursively seek the "in-between" shapes of consecutive models in the corresponding shape space. To respect the physical rationales, we embed the shape space into a reduced modal space spanned by deformation modes of the sample shapes. Meanwhile, to facilitate high-quality applications, we employ a geometric-modal hybrid scheme to separately handle dominant shape deformation and local details. In particular, the salient contributions of this paper can be summarized as follows:

- We pioneer a novel physics-plausible 4D shape subdivision scheme by encoding the potential deformation priors into shape interpolation in reduced modal space, which gives rise to interactive and flexible shape control.
- We formulate the recursive intermediate shape subdivision as a global optimization problem of finding a minimal energy-consumption path throughout the sparse samples, which can be solved by an Gaussian-Newton solver.
- We couple local geometry-space shape interpolation with global modal-space subdivision, which can simultaneously guarantee high-quality "in-between" shapes’ generation and its interactive efficiency.

### 2. Modal Space Definition and Construction

Our modal space definition and construction involve two main numerical techniques: modal analysis and deformation modes derivation. In the followings, we assume a mesh $\mathcal{M}$ is represented by a long vector $x \in \mathbb{R}^{|V|}$ comprising all vertex coordinates. Given a few sparsely-sampled models $\bar{x}_1, \bar{x}_2, ..., \bar{x}_k \in V$ along the temporal axis, our method produces a dense shape sequence which gives rise to natural transition from one key-frame to its succeeding one.

#### 2.1. Modal Analysis

For physics-based deformable objects and their dynamic simulation, the displacements $u = x - \bar{x} \in \mathbb{R}^{|V|}$ due to small deformations follow the equation

$$M\ddot{u} + D\dot{u} + Ku = f,$$

where $M, D, K \in \mathbb{R}^{|V| \times |V|}$ are the mass, damping, and stiffness matrix of the system respectively, and $f \in \mathbb{R}^{|V|}$ represents the external force. Specifically, $n$, the number of the vertices, is usually very large, which gives rise to efficiency bottleneck in practical applications.

**Linear Modal Analysis.** From the intrinsic view, elastic objects naturally have biased deformation directions, in which they deform with the least incremental deformation energy. For example, for a free-vibrating string, the low-frequency modes in nature correspond to the deformations that will consume minimal elastic strain energy. This also occurs on arbitrary 3D elastic objects. And such modes can be obtained by solving the generalized eigenvalue problem

$$Kx = \lambda Mx,$$

where the eigenvalues $\lambda_i$ represent the vibration frequencies, and the eigenvectors $\phi_i$ ($i = 1, 2, ..., k$) are vibration modes.

**Reduced Space.** Since the linear modes physically indicate the natural deformable directions of the objects, we can project the full-dimension shape onto a much smaller reduced space spanned by linear modes, and represent the shape via linear combination of these modes as

$$x = \bar{x} + \sum_{i=1}^{k} a_i \phi_i,$$

where $a = (a_1, a_2, ..., a_k)^T \in \mathbb{R}^k$ is a vector of modal amplitudes (typically $k \ll 3n$).
2.2. Modes Derivation from Deformation Energy

To embed the potential physical deformations in the modes, we should first compute the stiffness matrix $K$. As the way in [HSvTP10], it can be replaced by the Hessian of the mesh’s deformation energy.

Different from [HSvTP10], we define the energy on volumetric (tet) mesh. As shown in Fig. 3, the modes derived from this energy has the feature of volume preservation, which mainly benefits from the internal structures that can support the surface not shriveling up. Here we resort to as-rigid-as-possible (ARAP) deformation energy used in [H-WAG09], where they give an analytical expression of the Hessian of this energy, which makes it easy and efficient to compute the vibration modes.

After the linear modes are computed, we only capture the deformations on the surface. Because the modes with and without the points inside the surface are the same when they act as spatial bases, and it helps to gain efficiency if fewer points are taken into consideration in the optimization.

3. Modal Space Subdivision for Shape Completion

Based on the constructed modal space, in this section, we design our new global optimization based recursive subdivision framework for shape sequence completion.

3.1. Subdivision Scheme

Given a few sparse sampled points in Euclidean space, the most straightforward way to obtain a smooth curve that goes through all these points is to find the intermediate points between two consecutive ones while guaranteeing the bending energy of the curve is minimized.

In this paper, we extend the same subdivision idea to perform shape sequence completion in 4D space. Given a few sparsely-sampled models, each shape can be seen as a point in a high dimensional shape space, and thus finding a continuous shape sequence is equivalent to finding a smooth curve in shape space along which the integrated strain energy is minimized. Therefore, we should first focus on the mid-shape interpolation of two consecutive shapes.

3.2. Global Optimization for Mid-Shape Interpolation

Given two shapes $\mathbf{x}_1$ and $\mathbf{x}_2$, the desirable mid-shape $\mathbf{x}_{mid}$ should have minimal deformation energy w.r.t the two given ones. To increase the efficiency and respect the physical rationale, the aforementioned modal analysis process has to be applied first to reduce the space dimensionality and preserve physical meanings at the same time. However, there is a problem that must be addressed. Because there are two modal space $V_1 = \mathbf{x}_1 + \text{span} \{ \phi_1' \}$ and $V_2 = \mathbf{x}_2 + \text{span} \{ \phi_2' \}$, as shown in Fig. 4, it is extremely difficult to bridge them to represent the mid-shape under a unified framework. Thus, we need to solve this problem from another perspective. For a shape $\mathbf{x}_1 \in V_1$ and a shape $\mathbf{x}_2 \in V_2$, we enforce them to get close to each other in $\mathbb{R}^{2k}$, and make sure their deformation energies w.r.t their references are almost the same.

Objective Function Definition. As shown in Fig. 4, $d_2$ measures the "similarity" of $\mathbf{x}_1$ and $\mathbf{x}_2$, which forces them getting closer to each other in the optimization. And $d_1, d_3$ measure the "distance" between $\mathbf{x}_1$ and $\mathbf{x}_1, \mathbf{x}_2$ and $\mathbf{x}_2$ respectively, which help to preserve their initial shapes. We try to minimize the sum of $d_1, d_2, d_3$.

$$\min_{\mathbf{a}} d_1(\mathbf{x}_1, \mathbf{x}_1) + \mu d_2(\mathbf{x}_1, \mathbf{x}_2) + d_3(\mathbf{x}_2, \mathbf{x}_2),$$

where $\mathbf{a} = (\mathbf{a}_1^T, \mathbf{a}_2^T)^T \in \mathbb{R}^{2k}$ is a vector comprising two modal amplitudes of $\mathbf{x}_1$ and $\mathbf{x}_2$ (refer to Eq. 3), and $\mu > 0$ controls the balance between the "similarity" and "distance". It should be noted that, the solution space has been reduced from $\mathbb{R}^{2k}$ to $\mathbb{R}^{5n}$ (where $2k << 3n$) after modal analysis.

In Eq. 4, $d_1$ and $d_3$ are respectively the "distance" of $\mathbf{x}_1, \mathbf{x}_2$ and their reference meshes, just similar to discrete shells used in [HSvTP10], we measure this with stretching and bending terms, as well as a volume preservation term:

$$d_1(\mathbf{x}, \mathbf{x}) = \alpha \sum_{e \in \mathcal{E}} (\bar{L}_e - L_e)^2 \frac{1}{L_e^3} + \beta \sum_{e \in \mathcal{E}} (\bar{\Theta}_e - \Theta_e)^2 \frac{\bar{A}_e^2}{\bar{A}_e} + \gamma (\bar{V} - V)^2 \frac{1}{V^2}.$$

where $L_e, \bar{L}_e$ are the edge lengths in $\mathbf{x}, \bar{\mathbf{x}}$, $\Theta, \bar{\Theta}$ are the corresponding dihedral angles, and $\bar{V}, V$ are the volumes of $\mathbf{x}, \bar{\mathbf{x}}$. And $\bar{A}_e$ is the combined area of the two triangles sharing edge $e$ in $\bar{\mathbf{x}}$. while $d_3$ has the same format of $d_1$. The stiffness parameters above do not have to be adjusted for every model, and they are generally set to $\alpha = 500, \beta = 10$, and...
\( \gamma = 1000 \) in our experiments. As for \( d_2 \), we directly use the Euclidean distance to measure the similarity:

\[
d_2(\mathbf{x}_1, \mathbf{x}_2) = ||\mathbf{x}_1 - \mathbf{x}_2||^2. \tag{6}
\]

Finally, the objective function is defined in the reduced subspace \( \mathbb{R}^{2g} \). Generally speaking, it can well describe the potential deformation of a model with only 20-100 modes. And Eq. 4 is a standard nonlinear least-square problem, which can be solved by a Gaussian-Newton solver efficiently. Besides, the small dimension of the final solution makes it possible for interactive applications.

**Mesh Simplification Assisted Approximation.** Directly benefitting from the model reduction, the overall computational efficiency has been greatly improved. Nevertheless, the evaluations of the energy and computation of the Jacob matrix in G-N iterations still render quite a significant computational cost, and this is especially true when the shape has too many vertices and edges. To make these tasks independent of shape complexity, we need to continue to accelerate the energy and gradient computation by other strategies.

Similar to the way in [HSTP11], we also use an edge-collapse scheme to generate the coarse mesh with around 1000 faces and 500 vertices (actual numbers depend on mesh complexity), and each mesh is simplified with exactly the same rule to keep the topology consistent across the coarsened meshes. Edge-collapse schemes generate a mapping from the vertices of the fine original mesh to the vertices of the fine mesh that has one corresponding vertex on the collapsed coarse mesh. Let \( v' \) be a vertex of the coarse mesh and let \( \{v_1, v_2, \ldots, v_N\} \) be the set of vertices on the fine mesh that are collapsed to \( v' \), we compute the simplified modes via \( \phi'(v') = \frac{1}{N} \sum \phi(v_k). \) After that, we can conduct G-N iterations on the coarsened meshes. Finally, we use the obtained modal amplitudes \( \mathbf{a} \) together with the original modes to compute the shapes.

**Local Geometry Interpolation.** So far, modal analysis helps us reduce the shape space down to an extremely smaller modal space and the interpolation problem can be efficiently solved to achieve an interactive rate. However, the model reduction also brings one problem that geometric details are ignored. In the interpolation case, this manifests as the deformations in local regions are not well-interpolated, as shown in the top row of Fig. 5(b).

Therefore, we have to deal with local deformations separately. First, we should find which region should be treated distinctively. In [HWAG09], Huang et al. defined local fitting error based on the given deformation modes and decomposed the shape into different rigid parts. According to our observation, the points with highest local fitting error often lie on the shape extremities as shown in Fig. 5(a). And these are exactly the regions whose deformations cannot be covered well by linear modes. So, we employ Huang’s method to find the local regions and segment them apart. After that, when the energy minimization above is done, we further treat \( \mathbf{x}_1^* \) and \( \mathbf{x}_2^* \) as reference shapes, and interpolate the local regions to get new \( \mathbf{x}_1^* \) and \( \mathbf{x}_2^* \) with details, as shown in the bottom row of Fig. 5(b). In general, these local parts only consist of a few vertices, many existing methods can be used here. For example, we use the example-based method in [FB11] which is very relevant to our formulation.

**Median-Shape Generation.** When energy minimization is obtained, we can get two new shapes, \( \mathbf{x}_1^* \) and \( \mathbf{x}_2^* \). Since these two shapes are very close to each other in shape space, we can resort to a linear interpolation to get the final mid-shape via \( \mathbf{x}_{mid} = \frac{1}{2} (\mathbf{x}_1^* + \mathbf{x}_2^*) \).

### 3.3 Recursive Subdivision for Shape Completion

When we get one mid-shape, we can recursively compute more “in-between” shapes by taking them as new reference shapes. However, the modal analysis process is only conducted in the initial stage, it does not allow us to dynamically update modal modes after producing new shapes. So
we reuse the modes of the original mesh, which is closest to the newly-inserted shape. After 2 or 3 level recursions, the obtained “in-between” shapes are adequate to produce a smooth transition via linear interpolation, as shown in Fig. 6.

Figure 6: The subdivision results for two hand poses. They are almost the same as those produced by [KMP07].

4. Experimental Results and Evaluations

We have implemented our 4D shape sequence completion method using C/C++ on a PC with Intel Core i7-3370 CPU@ 3.40GHz. To verify the effectiveness and versatility, we have conducted different kinds of experiments. And the time performances of our experiments are detailed in Table 1. The table columns from left to right respectively represent: number of mesh vertices (#V), number of tetrahedral mesh vertices (#N), number of key-frames (#K), dimensions of the reduced space (#D), initialization time cost (IT), and average time cost for each interactive update (AT).

In all the experiments, to facilitate the numerical calculation of spectrum analysis, we have scaled the input shapes to an appropriate size. The parameters in Eq. 5 are set to constant values as described in Section 3.2. Therefore, we only interactively tune the parameters $\mu$ in Eq. 4 for different experiments to guarantee the result quality.

Fig. 1 shows a galloping horse sequence, which is interpolated from 6 key-frame models. Here only two of the input shapes and the refined “in-between” shapes are shown. Our result is shown in the top row, and the bottom row is the result produced by [HSvTP12]. There are obvious artifacts in [HSvTP12]’s results, for example, the horse’s right front hoof vanishes, where they are supposed to bend. In sharp contrast, our result can produce much better results. And this is totally because of the local geometry interpolation on the hooves, which enhance the details of the results. Moreover, this subdivision method also affords us the ability of data compression. In this experiment, using only 6 key-models can reconstruct the whole animation of the galloping horse, which helps reduce much storage to save key-models.

Fig. 2 and Fig. 6 respectively shows the experimental results of our modal subdivision method over woman models and hand models. Even if the deformation is large, our method can produce very natural transition from the source mesh to the target mesh. Without the guiding median shape, our result is also visually pleasing because of the fact that the prior knowledge is already implicitly encoded in the modal analysis. Moreover, the subdivision results are as good as the those from [KMP07], which conducts geodesic interpolation based on the as-isometric-as-possible shape metric.

Figure 7: The shape sequence interpolation comparison between the example-based method in [FB11] (right) and our method (left) over two given flamingo poses.

Fig. 7 compares the example-based method [FB11] with our modal subdivision method over two given flamingo poses. In [FB11]’s result, the left leg of the flamingo concaves to a flat shape and the “in-between” shapes deform in wrong direction. Because the leg is extremely slender, the example-based method can not guarantee the shape not sinking even with an volume preserving term. However, our deformation modes derived from volumetric mesh impose restrictions on the deformation, which can ensure the quality.

In addition, our method also has the ability to conduct cross-shape morphing. Fig. 8 respectively shows the shape morphing result of two lion poses and a cross-shape morphing between a lion and a cat. We firstly make the two meshes have the same topology by registering them with the blended intrinsic maps method [KLF11], and then re-mesh one of them according to the other’s connectivity. After that, the following steps are the same as the aforementioned subdivision process. Even though this deformation is not very large, but it is nontrivial to deal with the tail deformation correctly. As demonstrated in Fig. 10 of [GLHH13], the results from [HAWG08] have visible artifacts, wherein the tail either twists unnaturally or becomes shorter.

Figure 8: The shape morphing results: between two lion poses (top), and between a cat model and a lion (bottom).

Performance and Accuracy. In general, the dimension of
the reduced shape space is very small, for example, \(2k = 50 + 2 = 100\), which guarantees Eq. 4 being efficiently solved. In the experiments above, our solver can commonly find the minimum within 5 iterations. In practice, we uniformly set the max iteration number to be 10, so that we can achieve the accuracy of \(10^{-10}\) during energy decrease.

Table 1 shows the timing statistics of all experiments documented above. Despite the initial stage, even with complex models (more than 26K vertices), our framework responds very quickly. This is mainly because of the modal analysis and gradient approximation process, which makes the performance independent of the complexity of the input models. It should be noted that, extra local interpolations are added for the Flamingo, Horse, and Hand models and they need a little more time in the update stage.

Table 1: Performance statistics (in seconds)

<table>
<thead>
<tr>
<th>Model</th>
<th>#V</th>
<th>#N</th>
<th>#K</th>
<th>#D</th>
<th>IT</th>
<th>UT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lion</td>
<td>5000</td>
<td>29.5K</td>
<td>2</td>
<td>60</td>
<td>39.10</td>
<td>0.09</td>
</tr>
<tr>
<td>Cat</td>
<td>7207</td>
<td>41.4K</td>
<td>2</td>
<td>60</td>
<td>58.36</td>
<td>0.09</td>
</tr>
<tr>
<td>Flamingo</td>
<td>26284</td>
<td>103K</td>
<td>2</td>
<td>100</td>
<td>153.01</td>
<td>0.73</td>
</tr>
<tr>
<td>Horse</td>
<td>8431</td>
<td>51.4K</td>
<td>6</td>
<td>100</td>
<td>69.24</td>
<td>0.32</td>
</tr>
<tr>
<td>Hand</td>
<td>13071</td>
<td>50.2K</td>
<td>2</td>
<td>100</td>
<td>31.86</td>
<td>0.28</td>
</tr>
<tr>
<td>Women</td>
<td>25381</td>
<td>93.7K</td>
<td>2</td>
<td>100</td>
<td>96.06</td>
<td>0.12</td>
</tr>
</tbody>
</table>

5. Discussion and Conclusion

In this paper we have articulated a new hybrid shape interpolation method that can take advantages of both data-driven and physics-based approaches at the same time. Our novel idea is to transplant the recursive subdivision from high-dimensional curve generation to 4D shape completion from sparse model samples in reduced modal space. Our hybrid shape completion methodology and its system framework comprise several key technical components, including: vibration mode extraction based on volumetric deformation energy, modal analysis driven model reduction, subdivision scheme to recursively find “in-between” shapes, and a hierarchical and hybrid strategy to further improve the efficiency and guarantee deformation details. Besides, we have also conducted diverse types of experiments over sparsely-sampled key-frame 3D models, which all demonstrate the advantages and great potentials of our framework.

Despite the attractive properties of this methodology detailed in our system framework, our hybrid approach still has some limitations that should be overcome in our future work. The linearization of the underlying dynamic ODEs could not completely prevent undesirable shape artifacts, especially when dealing with large-scale deformations in the existence of too few model samples. Moreover, the average median-shape of \(x_1\), \(x_2\) may be unstable when the model has a very large hollow body.

6. Acknowledgements

This research is supported in part by National Natural Science Foundation of China (Nos. 61190120, 61190121, 61190125, and 61300067) and National Science Foundation of USA (Nos. IIS-0949467, IIS-1047715, and IIS-1049448). The authors would like to thank the anonymous reviewers for their helpful comments.

References


(c) The Eurographics Association 2015.