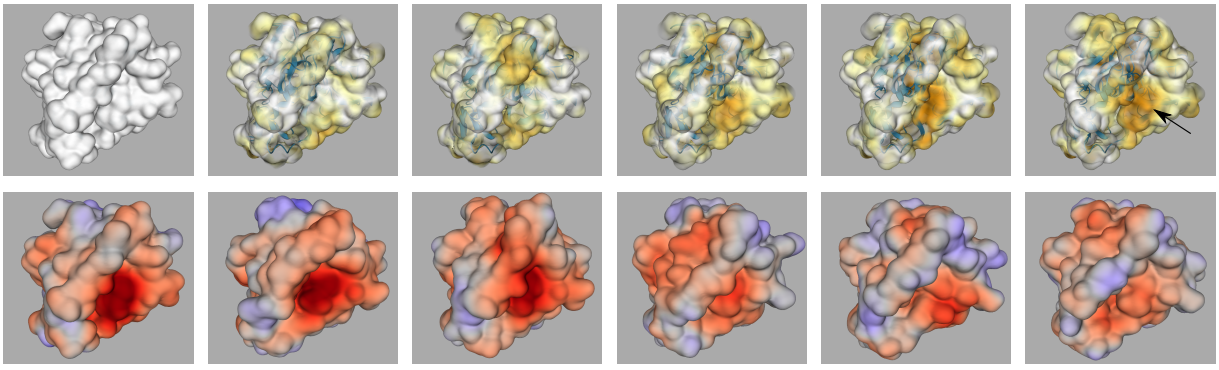
DOI: 10.1111/cgf.12375
Eurographics Conference on Visualization (EuroVis) 2014
H. Carr, P. Rheingans, and H. Schumann
(Guest Editors)

Volume 33 (2014), Number 3

Comparative Visualization of Molecular Surfaces Using Deformable Models

K. Scharnowski[1], M. Krone[1], G. Reina[1], T. Kulschewski[2], J. Pleiss[2], and T. Ertl[1]

[1]Visualization Research Center, University of Stuttgart, Germany
[2]Institute of Technical Biochemistry, University of Stuttgart, Germany



(a) MeOH = 0.00 *(b) MeOH = 0.01* *(c) MeOH = 0.16* *(d) MeOH = 0.46* *(e) MeOH = 0.73* *(f) MeOH = 0.77*

Figure 1: *Electrostatic potential (bottom row) and comparative visualization (top row) of six C. antarctica Lipase B (CALB) simulation results with increasing methanol (MeOH) activity. The top row shows a comparison of conformer one (a) with respect to all other conformers (b-f). The gradual change in the electrostatic potential complies with the change in MeOH activity. The conformer shown in (a) differs most from the one in (f), which is shown in our rendering by a more saturated patch (arrow).*

Abstract

The comparison of molecular surface attributes is of interest for computer aided drug design and the analysis of biochemical simulations. Due to the non-rigid nature of molecular surfaces, partial shape matching is feasible for mapping two surfaces onto each other. We present a novel technique to obtain a mapping relation between two surfaces using a deformable model approach. This relation is used for pair-wise comparison of local surface attributes (e.g. electrostatic potential). We combine the difference value as well as the comparability as derived from the local matching quality in a 3D molecular visualization by mapping them to color. A 2D matrix shows the global dissimilarity in an overview of different data sets in an ensemble. We apply our visualizations to simulation results provided by collaborators from the field of biochemistry to evaluate the effectiveness of our results.

Categories and Subject Descriptors (according to ACM CCS): I.3.5 [Computer Graphics]: Computational Geometry and Object Modeling—Boundary representations, I.4.7 [Image Processing and Computer Vision]: Feature Measurement—Size and Shape, J.3 [Computer Applications]: Life and Medical Sciences—Biology and Genetics

1. Introduction

Analyzing molecular characteristics often involves comparing ensembles of similar molecules concerning their differences in function or structure. This scenario is common in drug design, where the function of proteins is often investigated by executing Molecular Dynamics (MD) simulations under different conditions (e.g. in varying solvent mixtures) or with selective mutations of amino acids. Further scenarios are the comparison of proteins from different organisms that presumably have the same function or the evaluation of simulation results across multiple scales. The ensemble

members are often analyzed by grouping them based on the principle of molecular similarity, which states that molecules with similar geometric structure and fold are likely to exhibit similar functions [JM90]. Since the molecular surface is the interface between the molecule and its surroundings, its characteristics are essential for function. If, for example, a ligand binds to the molecule, it needs to fit geometrically and the surface properties have to be complementary so that it is not repelled (e.g. electrostatic surface potential). Hence, the shape and attributes of two molecular surfaces can be compared to analyze the change in functionality [VFBHC00]. To this end, automated screening of data bases using shape matching is a common analysis approach [FM13], in particular in medical image analysis and biochemistry. A global measurement helps to get an overview over the members of an ensemble, whereas for a detailed comparison of two molecules concerning differences in their function, a local measure of similar and dissimilar surface areas is beneficial.

The problem of establishing a meaningful relation between two or more shapes is well known and there has been a rich history of studies in various fields of research, such as image analysis [SCP12], computer graphics [vKZHCO11], and shape matching [TV08]. In this work, we present a comparative visualization method for 3D surfaces that uses vertex tracking to establish a correspondence between two input surfaces. Special attention is given to biomolecular surfaces, whose characteristic behavior imposes several additional challenges. First, especially in the context of biochemistry, only parts of the input surfaces might be comparable in a meaningful way and, hence, only partial shape matching might be applicable. This is due to the global deformation of the molecular structure and the possible genus changes of the molecular surface that come with the deformation. Second, molecular surfaces exhibit a high frequency of concave and convex regions. Explicit vertex tracking on such surfaces can lead to unwanted intersection of surface parts in their final position. In our approach, we overcome this difficulty by using a deformation process that takes the surface normal of the source surface and the target one into account. Finally, the matching algorithm should work without additional user input to make it applicable to the screening of larger data bases. Hence, a meaningful correspondence has to be established using purely data-driven features that require no user input.

We present a method that allows comparing arbitrary attributes of two input molecular surfaces. To this end, we use a deformable model approach for non-rigid partial shape matching. A source surface is triangulated and deformed to match the target using a velocity field that is based on a diffusion process. Consequently, a vertex-wise correspondence between the source surface and the target surface is established. We compute a dissimilarity measurement for each pair of points and associate the result with a quantification of the surface comparability. Our method uses both a local and a global dissimilarity descriptor to provide an overview as well as a detailed comparison. The global dissimilarity is depicted in a 2D matrix plot to allow for the analysis of large numbers of pair-wise comparisons. After identifying an interesting pair of input data sets, a comparative 3D visualization of the two respective surfaces can be shown for detailed analysis. This 3D visualization shows the local dissimilarity attribute and the geometrical differences color-coded on the matched surface regions of the target surface. We evaluate our technique by applying it to different real world ensembles of MD simulation results provided by domain experts working in the field of biochemistry. Our main contribution is a novel technique for comparative visualization of molecular surfaces that uses a deformation process to obtain both a local and global dissimilarity measurement for arbitrary surface attributes.

2. Related Work

The comparison of molecules based on geometry is a common approach in computational chemistry. Chimera [PGH*04] offers a structure-based morphing animation between two input molecules by interpolating the atom positions. However, there is no numerical comparison of surface geometry or surface attributes. In the field of computational chemistry, there are a number of surface-based matching algorithms, which use either global or local similarity measurements. Global measurements are distance functions based on 3D Zernike descriptors [SLL*08] or based on a parametrization of the surface with spherical harmonics [RK99]. A deformation invariant shape descriptor is presented in [LFR09]. Local similarity measurements compare individual patches [HLA04] or consider local surface curvature [CBJ00]. Other work projects surface properties to an intermediate domain such as a sphere [PS09]. In [BH06], evenly distributed surface vertices are matched using an iterative closest point scheme to establish correspondence. Our approach uses a deformation process to track points and combines the local deformation and physico-chemical properties of the surface in a comparative visualization.

Shape correspondence problems play an important role in various fields of research related to geometry processing [MDA01] and image analysis [SDP13]. Early work in these fields uses deformable models for image segmentation [KWT88, TWK88]. Deformable models represent geometric shape either implicitly by a level-set or explicitly by a discrete mesh. Models with explicit representation can easily be used to create a one-to-one mapping between genus zero input shapes. In [SAPH04], a coarse patch layout is built in both input surfaces with subsequent re-meshing to establish correspondence. In [KS04], an explicit cross-parametrization between two surfaces is computed. However, explicit models do not naturally handle objects of arbitrary genus. For implicitly represented shapes, the deformation is executed by interpolating the source and target scalar

functions [OF03]. Here, genus changes are handled naturally, but explicit point tracking requires additional work.

The problem of establishing shape correspondence between surfaces of arbitrary genus while being able to track points is an actively researched field. In [DYT05], explicit point tracking is performed based on a prior implicit morph. Other work suggests morphing between objects of arbitrary genus by combining distance field interpolation with point tracking [WCX*13]. Here, point-to-point correspondence requires user input in the form of sparse anchor points. Other methods repeatedly resample the implicit surface for tracking [BSM*06, Mül09]. Bojsen-Hansen et al. [BHLW12] represent surfaces explicitly, whereas for handling different topology, they are voxelized and resampled. Newly created points are tracked in both directions to preserve correspondence. In our work, we also track newly created points backwards to achieve a more consistent sampling of surface attributes. There are also mesh-based surface tracking methods that use velocity fields to evolve a given triangle mesh. [WTGT10] propose a mesh-based surface tracking method for fluid animation. [CK10] present a technique to implement operators that modify the topology of polygonal meshes at intersections and self-intersections.

The aforementioned methods aim at creating a full bijective mapping while minimizing overall texture distortion. In contrast, we are aiming at a partial mapping that reduces distortion in surface regions that we consider comparable to a minimum while allowing higher distortion at locations that do not yield a meaningful comparison. Additionally, we want the deformation process to be symmetric to improve the quality of our global descriptor. We can, therefore, use a deformable model approach in order to establish shape correspondence on the vertex level. Such a method distinguishes between external and internal forces that act on the model and gives us greater control over the way the input mesh deforms. The *Gradient Vector Flow (GVF)*, introduced by Xu et al. [XP98], is an external force that remedies some problems of the earlier deformable model approaches by applying diffusion to an initial velocity field. In our approach, we incorporate the surface normals of both input surfaces in the GVF computation to achieve a symmetrical transition between the input surfaces. There exist extensions to the original GVF, which take directional information of the target shape into account [TMA*06]. Here, additional user input is used to specify the orientation of the target shape. In our case, the orientation of the target is given by the input data. We can, therefore, use a simplified variation of the original GVF with different initial forces for the diffusion process, which incorporate the surface orientation of both surfaces in the deformation process.

3. A Mapping Relation for Molecular Surfaces

Molecular surfaces that have been relaxed by an MD simulation often undergo local deformations that change the surface geometry. In general, there are areas on the input surfaces that we do not consider meaningfully comparable. These include areas where there are strong local deformations, strong global deformations, and genus changes. In fact, partial shape matching is a method commonly used in computational drug design [SB08] to compare molecules. We, therefore, aim at a partial mapping $m : S \rightarrow T$ between the source surface S and the target surface T that leads to low distortion in areas that we consider comparable.

We approximate the *solvent excluded surface (SES)* [Ric77] using *Meta Balls* [Bli82]. This approximation has been shown to be sufficiently accurate for analysis [GP95]. Each particle of the molecule is associated with a Gaussian density distribution. Subsequently, a volume is computed by accumulating the density contributions of all particles at each voxel, which yields smooth density maps V_S and V_T.

Before S is locally deformed, rigid alignment is used to spatially superimpose the two input surfaces. For arbitrary surfaces, purely geometrical rigid alignment algorithms like *Iterative Closest Points* [BM92] could be applied. In our application scenario, however, a meaningful alignment can easily be obtained by using *root-mean-square deviation (RMSD)* [Kab76] minimization of the atomic positions. RMSD minimization is common practice in bioinformatics and is less sensitive to symmetries of the surfaces then purely geometrical approaches. For the RMSD alignment of very different structures, an appropriate numbering scheme needs to be provided that associates amino acids of both input structures, however, using the backbone atoms is sufficient in most cases. By aligning the molecular structures, the position and orientation of the implicitly defined surfaces is optimized with respect to structural similarities.

To address the challenges regarding molecular surface mapping, we use a combination of rigid and non-rigid alignment. Using a deformable model approach for non-rigid alignment is a natural choice, since it allows controlling and explicitly tracking the movement of the surface points. After the deformation step, the start and end positions of the vertices of the source mesh define a trivial correspondence between the two surfaces. We use the mapping relation to define two kinds of local descriptors. The first local descriptor d is computed by sampling a specific surface attribute at pairwise corresponding surface points and computing a difference measure based on the obtained values. The second descriptor g quantifies the local differences in surface geometry. We use the second descriptor as an indicator for the comparability of the local surface attributes, since higher variation in surface geometry indicates worse comparability of the surfaces areas.

4. Non-rigid Shape Propagation

The deformation is performed on a mesh that has been extracted from the aligned source structure. Using a de-

formable model approach allows us to represent the molecular surface as an elastic shape—similar to a spring mass model—that can be deformed towards the target shape. We use a deformable model approach that allows us to define the deformation forces in a way that aims at generating low local distortion in well-comparable areas of the input surfaces.

4.1. Deformable Models

A deformable model or—in this case—a deformable surface is an elastic shape whose deformation is described by internal and external energy, where the internal energy term stands for the material properties of the surface (e.g. tension, rigidity) and the external energy is defined to pull the surface to its target shape. Deformable models were first introduced by Kass et al. [KWT88] and were later generalized to the 3D case [TWK88]. The objective of the deformable model method is to minimize an energy functional

$$E(\mathbf{s}) = E_{int}(\mathbf{s}) + E_{ext}(\mathbf{s}) \rightarrow 0, \tag{1}$$

where $\mathbf{s}$ is a point on the elastic surface S, given by its 3D coordinates. The vertex displacements are obtained by computing a force equilibrium between external and internal forces [CC93]:

$$\mathbf{F}_{int} = -\underbrace{\tau\Delta\mathbf{s}}_{tension} + \underbrace{\rho\Delta^2\mathbf{s}}_{rigidity} = -\mathbf{F}_{ext}, \tag{2}$$

where Δ is the Laplacian. τ and ρ control tension (rubber skin behavior) and rigidity (thin plate behavior) of the surface.

Using the dynamic force-based formulation, the evolution of the deformable surface can be computed iteratively in discrete time steps, until it becomes stationary [CC93]. The position of a surface point $\mathbf{s}$ at time step $t+1$ is

$$\mathbf{s}^{(t+1)} = \mathbf{s}^{(t)} + \sigma((1.0-\mu)\,\mathbf{F}_{int}(\mathbf{s}) + \mu\,\mathbf{F}_{ext}(\mathbf{s})), \tag{3}$$

where μ is used to steer the influence of internal and external forces and σ is a scaling factor to ensure convergence.

4.2. Definition of the External Forces $\mathbf{F}_{ext}$

We want to define an external force that pulls similar regions of the input surfaces onto each other, while leaving out areas where a point-to-point mapping would not be meaningful. Here, several requirements have to be met. Often, differently sized surface patches correspond to each other. Consequently, the triangles in the respective surface patch of the source mesh are rescaled during the deformation process. We want this rescaling to happen in a manner that preserves the overall vertex distribution in the source mesh. This can be achieved by defining an external force that leads to a evenly spaced distribution of vertex paths.

The external forces proposed in the original deformable model approach [KWT88] are obtained by using the gradient of the input images edge map. These forces have a limited capture range and they often fail to move source points into highly concave regions. An external force called *Gradient Vector Flow (GVF)* proposed by Xu et al. [XP98] addresses these two problems. The GVF has a potentially infinite capture range and properly handles concave regions. It is computed by initializing the external forces with the image-based gradient near the target edges. If the initial external force is $\tilde{\mathbf{F}}_{ext}$, then the GVF is defined as the vector field $\mathbf{v} = (u, v, w)$ that fulfills the following Euler equation (see [XP98] for details):

$$r\Delta\mathbf{v} - \|\tilde{\mathbf{F}}_{ext}\|^2(\mathbf{v} - \tilde{\mathbf{F}}_{ext}) = 0. \tag{4}$$

Here, r is a smoothing factor to steer the influence of the diffusion (needed for noisy input data). Intuitively, solving Equation 4 models a diffusion process. To account for the low capture range of $\tilde{\mathbf{F}}_{ext}$, equation 4 aims at using $\tilde{\mathbf{F}}_{ext}$, if the magnitude is sufficiently high and using the diffusion otherwise. Consequently, $\mathbf{v}$ spreads smoothly through the volume, starting with a direction that complies with the initial external forces $\tilde{\mathbf{F}}_{ext}$.

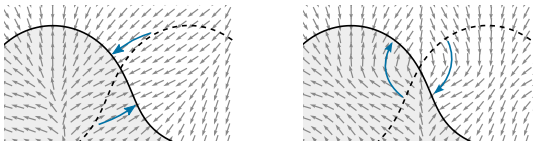
In our approach, we modify the GVF to create a symmetric transition since an asymmetric deformation process would ultimately lead to an asymmetric outcome of our global difference measurements. We initialize $\tilde{\mathbf{F}}_{ext}$ with the normalized gradients of the Gaussian volumes V_S and V_T at the border regions of both the target and the source surface, respectively. In a general case, the surface normals could be sampled on a grid to be used as initial external force $\tilde{\mathbf{F}}_{ext}$. If both surfaces overlap, we use the average of the two gradients as the initial value. We do not need the smoothing factor r, since our input data is noise-free. Additionally, we also want the exact values at the border regions to stay fixed. We, therefore, set $r = (1 - \|\tilde{\mathbf{F}}_{ext}\|^2)$.

After calculating the diffusion using Equation 4, the vector field orientation of $\mathbf{v}$ is parallel to the normals of both the source and the target surface, connected by a smooth transition. Since the volume gradient computed from the Meta Balls representation always points towards atom centers, we subsequently invert $\mathbf{v}$ at locations that lie inside the target volume. Consequently, the resulting field $\mathbf{v}$ pulls the source mesh towards the target surface (Figure 2b).

Equation 4 is solved by treating the components of the GVF field as a function of time and reformulate the problem using discretization in space and time (see [XP98] for more details). The solution can then be found by solving three decoupled equations iteratively.

$$\begin{aligned} u_{i,j,k}^{n+1} &= c_{i,j,k}^1 + (1 - b_{i,j,k})(u_{i,j,k}^n + d_t(\Delta u_{i,j,k})) \\ v_{i,j,k}^{n+1} &= c_{i,j,k}^2 + (1 - b_{i,j,k})(v_{i,j,k}^n + d_t(\Delta v_{i,j,k})) \\ w_{i,j,k}^{n+1} &= c_{i,j,k}^3 + (1 - b_{i,j,k})(w_{i,j,k}^n + d_t(\Delta w_{i,j,k})) \end{aligned} \tag{5}$$

where $b = f_x^2 + f_y^2 + f_z^2$, $c^1 = bf_x$, $c^2 = bf_y$, and $c^3 = bf_z$ and



(a) Original GVF (b) Modified GVF

Figure 2: *Comparison of the original GVF [XP98] and our modified GVF. In the regions marked by the blue arrows, the original GVF causes the source surface S (dashed line) to be pulled towards the wrong side of the target surface T (solid line). This would lead to self-intersection in the deformation process. This issue is resolved when using our modified GVF.*

the Laplacian for the vector components is approximated as

$$\begin{aligned}\Delta u_{i,j,k} =&(u_{i+1,j,k}+u_{i-1,j,k}+u_{i,j+1,k}+\\ &u_{i,j-1,k}+u_{i,j,k+1}+u_{i,j,k-1}-6u_{i,j,k})\\ \Delta v_{i,j,k} =&(v_{i+1,j,k}+v_{i-1,j,k}+v_{i,j+1,k}+\\ &v_{i,j-1,k}+v_{i,j,k+1}+v_{i,j,k-1}-6v_{i,j,k})\\ \Delta w_{i,j,k} =&(w_{i+1,j,k}+w_{i-1,j,k}+w_{i,j+1,k}+\\ &w_{i,j-1,k}+w_{i,j,k+1}+w_{i,j,k-1}-6w_{i,j,k}).\end{aligned} \tag{6}$$

The forces obtained by our approach exhibit different behavior than the original GVF forces. The path of the vertices starts orthogonally to the source surface and terminates orthogonally at the border of the target surface. This prevents self-intersection in cases where two convex parts overlap (Figure 2). We do not use internal forces in the mapping deformation since they would introduce additional distortion.

5. Local and Global Difference Measurements

The mapping relation established by the algorithms described before can be used to compare arbitrary surface attributes of both input shapes. We place special attention on the electrostatic surface potential since it is an important indicator for the function of biomolecules. We define two difference measurements taking both the absolute value of the potential and the potential sign into account. A potential difference d_ϕ is computed as

$$d_\phi(\mathbf{s}) = |\phi(\mathbf{s}) - \phi(m(\mathbf{s}))|, \tag{7}$$

where m is the mapping relation and ϕ is the sampled electrostatic potential at the surface point $\mathbf{s}$. The potential sign measurement d_{sign} is defined by

$$d_{sign}(\mathbf{s}) = \begin{cases} 1, & \text{if } \operatorname{sign}(\phi(\mathbf{s})) \neq \operatorname{sign}(\phi(m(\mathbf{s}))) \\ 0, & \text{otherwise} \end{cases}. \tag{8}$$

Due to local geometrical differences of the surfaces, not all matched surface regions are equally well comparable. We quantify these geometrical differences the local deformation g, which is obtained for a vertex s by the accumulating the length of the vertex path during the iteration.

The local measurements d_ϕ and d_{sign} as well as the deformation factor g can be used to compute global descriptors by integrating their value over the surface area and then normalizing them by dividing the result by the total surface area. The global mean differences thus are

$$\begin{aligned} D_\phi(S,T) &= \frac{1}{|S|}\int_\Omega d_\phi(\mathbf{s})dS \\ D_{sign}(S,T) &= \frac{1}{|S|}\int_\Omega d_{sign}(\mathbf{s})dS \\ G(S,T) &= \frac{1}{|S|}\int_\Omega g(\mathbf{s})dS \end{aligned} \tag{9}$$

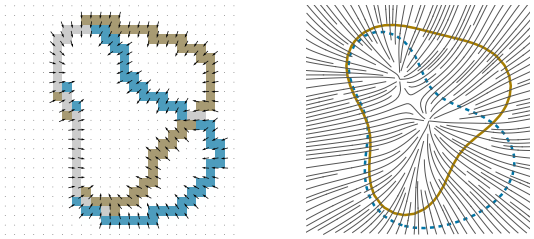
$D_\phi(S,T)$ models a mean error of the absolute potential. $D_{sign}(S,T)$ describes the percentage of the target surface T that has a different sign then the source surface S. $G(S,T)$ is a global descriptor for geometrical differences.

6. Implementation & Rendering

All computations, except the RMSD minimization, were parallelized on the GPU using CUDA. This includes volume creation, triangulation, computation of the GVF, mesh deformation including mesh refinement, and finally the difference measurements. The renderings were obtained using OpenGL/GLSL.

We extract the Gaussian volume using the parallel gathering approach proposed in [KSES12]. The surface mesh is extracted from the volume using a parallel implementation of the Marching Tetrahedra method. Our implementation is similar to the Marching Cubes one presented in [KSES12]. We then apply a regularization step to the mesh prior to the actual mapping deformation. This regularization step is necessary to obtain a mostly uniform vertex distribution and, hence, a uniform sampling of surface attributes of the source surface. To this end, we use the deformable model approach to relax the mesh on the level-set that defines it. For this step we use both internal and external forces. The external forces are based on the gradient of the Gaussian volume to keep the vertices positioned at the isovalue. The internal forces use $\tau = 0.5$ and $\rho = 0.5$, the force scaling σ is set to 1.0, and the external force weighting μ was 0.5. The required mesh Laplacian can be computed as described e.g. in [SHL*11].

In order to obtain the initial external forces f, the gradients of the Gaussian volume in the border region of both the source mesh and the target one are sampled using a parallel gathering approach. The gradient of a lattice point is sampled if any of its adjacent neighbor cells contain either the source surface or the target one (see Fig. 3a). If both surfaces are overlapping, the average of both gradients is computed and normalized, since this approximates the result after the diffusion of both gradient fields. Subsequently, equation 5 is solved iteratively for all three vector components at all grid points as described in section 4.2 until a stationary solution is found (see Figure 3b).



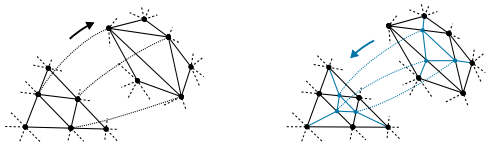
(a) Initial external forces *(b) Streamlines of final GVF*

Figure 3: *A discrete plot of the external forces computed with our GVF approach. (a) shows the initial external forces obtained from the gradients in the border regions of both the source surface S (blue) and the target surface T (ochre). (b) shows streamlines following the final external forces.*

Subsequently, the mesh representing the source surface S is deformed using the GVF forces. In our experiments, an initial value for the force scaling σ of $0.1 \times \Delta_g$ (with Δ_g being the grid spacing) yielded a good tradeoff between convergence speed and accuracy. For faster convergence the magnitude of the sampled external forces is divided by two every time the isosurface is crossed. To this end, we store the current force scaling for the individual vertices after each iteration step. We stop the overall iteration process if the average displacement of all vertices is smaller than the minimum displacement $\Delta_g \times 10^{-2} \times \sigma$. The final mesh is defined by the converged vertex positions while maintaining the topology of the source mesh.

In order to avoid additional mesh distortion in areas that are well comparable, we do not use internal forces during the deformation. The internal forces would cause the vertices to leave their path defined by the diffusion in order to regularize the triangles, thereby lowering the significance of the surface attributes. However, due to the discrete surface representation, this can lead to oversized unmapped triangles when a source region is associated with a target region of significantly bigger size (see Figure 4a). We consider a triangle unmapped if its vertices and their center of mass are not contained in a volume cell that is crossed by the target isosurface to account both for triangles spanning over cavities and triangles with non-converged vertex positions. In order to counteract asymmetrical behavior when computing D_ϕ, D_{sign}, and G, we apply adaptive mesh refinement after the deformation process. To this end, we first flag all edges that are longer then Δ_g and perform the subdivision for all triangles. We use a straight-forward subdivision-scheme yielding four, three, or two new triangles, respectively (see Figure 4b). The newly created vertices are then transported towards the target surface using to the GVF forces.

Finally, we compute the values of the global descriptors D_ϕ, D_{sign}, and G. g is obtained by tracking the vertex path and accumulating the discrete steps for each vertex. Since the newly created vertices stemming from the subdivision process do not have a properly tracked path we backtrack the deformation process for those vertices to obtain both the path length and their associated starting position (see Figure 4b). Based on the sampled electrostatic surface potential, we first compute d_ϕ and d_{sign} for all vertices. Subsequently, g, d_ϕ, and d_{sign} are integrated over the surface area of the morphed mesh and the results are normalized according to the target surface area. By default, unmapped triangles are omitted to account only for comparable surface parts.



(a) *(b)*

Figure 4: *(a) illustrates how the deformation can lead to oversized triangles. (b) shows our subdivision approach with subsequent backtracking for consistent sampling.*

The mapped surface parts are rendered semi-transparently, where g is mapped to opacity. The local difference measurement is color-coded using a HSV color mapping, where we vary the saturation accordingly (high saturation denotes high difference). We chose this simple approach over more sophisticated methods for uncertainty visualization (e.g. [PRW11]) to avoid visual clutter. For additional contextual clues, we combine the transparent surface rendering with two cartoon renderings representing the molecular structures. We also render unmapped triangles that might still be present—even after the subdivision—transparently, since they represent areas that do not have a comparable counterpart. The 2D matrix plot showing the global descriptors uses the same color map as the 3D rendering. The color map and the surface transparency are scaled according to user-defined thresholds. As a default, the minimum and maximum values of the local difference measure can be used. In order to provide an overview as well as to allow for detailed analysis, our tool allows the user to pick a pair-wise result in the 2D plot to show the respective 3D rendering in a separate window.

7. Results & Discussion

We tested our mapping pipeline on a system with an Intel Core i7-4770 with 16 GB RAM and a GeForce GTX 760 (2 GB VRAM). Computation times can be found in Table 1. The mesh refinement is the most time-consuming step since our un-optimized implementation requires us to reallocate memory. The diffusion process for the GVF calculation was carried out over 50 iterations. The regularization needed about 10 iterations on average until convergence. The mapping deformation process needed about 150 iterations on average until convergence. Our surface rendering is straightforward, hence, interactive frame rates can be maintained.

Table 1: *Performance of the surface mapping pipeline (all times in ms).* ***#Dat*** *ist the number of files in the ensemble,* ***#Vox*** *denotes the number of voxels of the density volume, and* ***#V*** *the vertex count of the source mesh. Computation times are given as* $\mathbf{t_{mesh}}$ *(initial mesh generation including alignment, volume generation, and mesh regularization),* $\mathbf{t_{gvf}}$ *(GVF),* $\mathbf{t_{map}}$ *(mapping deformation),* $\mathbf{t_{sub}}$ *(subdivision), and* $\mathbf{t_{meas}}$ *(global descriptors).* ***#Vox***, ***#V***, *and values in parentheses are ensemble means.*

	#Dat	#Vox	#V	t_{mesh}	t_{gvf}	t_{map}	t_{sub}	t_{meas}	$\sum t$
P450	24	617k	67k	40–67 (53)	50–81 (62)	3–45 (24)	50–1545 (407)	29–48 (33)	121–1705 (516)
TIM	26	402k	41k	29–62 (35)	34–62 (41)	2–26 (17)	47–197 (128)	28–43 (30)	106–328 (211)
CALB	152	435k	48k	37–79 (51)	35–63 (42)	2–125 (18)	42–285 (158)	28–44 (33)	108–596 (260)

The applicability of our method is evaluated together with collaborators from the field of biochemistry. Grouping and comparing of data sets is a time consuming task in their daily work. Our technique can be used to help them explore ensemble data based on the principle of molecular similarity. A possible work flow would be to first investigate the overview matrix to find either outliers or similar groups with respect to one reference data set. Subsequently, a 3D rendering of interesting data sets can be shown for more detailed analysis, e.g. to search for functional surface regions. We apply our approach to different ensembles obtained by MD simulations and synthetic data sets and discuss our results.

7.1. Local Dissimilarity

We tested our method with an ensemble of 24 variants of cytochrome P450 monooxigenase, which consist of varying amino acid sequences despite having a similar functionality. Here, our collaborators are interested in finding functional surface regions that remain conserved throughout the ensemble. Figure 5 shows a comparison of two examples (PDB IDs: 1S1F & 3N9Y), which both use a class I redox system. The cartoon rendering provides additional clues about the underlying molecular structure and helps to distinguish transparent parts from opaque parts of the surface. Several distinct regions are revealed where the absolute potential difference is high while the surface geometry is similar (arrow). Figure 5e depicts a region where a difference in surface geometry leads to a transparent patch in the comparative rendering. We also applied our method to 152 conformers (conformational variants) of *Candida antarctica* Lipase B (CALB). The structures were equilibrated in varying solvent mixtures (methanol, water and toluene) using MD simulations. We created 3D renderings for a subset with gradually increasing methanol activity and constant water activity in the solvent mixture to investigate the methanol's influence on the electrostatic surface potential. Figure 1 shows the results for this subset. The conformer at a methanol activity of 0.00 (1a) differs most from the one at the highest methanol activity of 0.77 (1f), which is indicated by a higher saturation. That is, our results suggest a change in the surface potential for increasing methanol activity, as expected.

In our deformation process we are aiming at little mesh distortion in surface areas that yield a meaningful comparison. Figure 6 shows the result when using our method to create a morph between a torus and a sphere. Figure 6b shows that there is higher mesh distortion in areas with worse comparability (hence a high value of g) and unmapped areas. In contrast, mesh distortion is lower in well comparable parts.

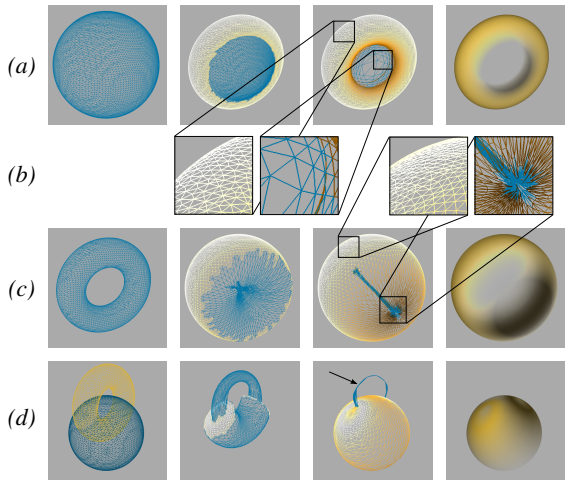


Figure 6: *Results for surfaces with different genus. The first column shows the source surface (blue) and the target surface (yellow). The second and third columns show intermediate and final mapping results (g is mapped to saturation; unmapped triangles are blue). The last column shows the result (with g mapped to opacity). (a) mapping a sphere to a torus. (b) high mesh distortion in unmapped and badly comparable areas. (c) mapping a torus to a sphere. (d) non-mappable areas collapse to a cluster of unmapped triangles when mapping a torus to a sphere with a positional offset.*

We do not apply special treatment to surfaces with different genus since our algorithm treats those cases by omitting unmapped surface areas. Figures 6a, 6c, 6d, and 7 show how unmapped regions develop in different scenarios. In Figure 6a, the outer parts of the torus are mapped while the middle part is omitted. In Figure 6c, stretched unmapped triangles appear inside the morphed shape. In Figure 6d, the non-comparable parts of the torus' outer arc collapse to a cluster of unmapped triangles. Figure 7 shows a scenario with molecular input data. Here, the source mesh is pulled into

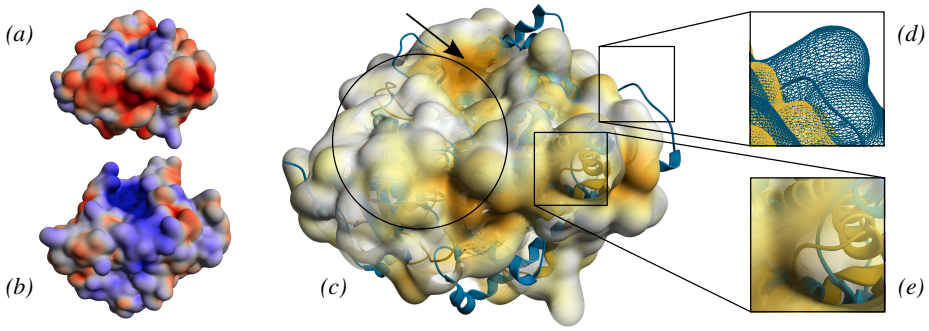


Figure 5: *A comparative surface rendering of two variants of P450. On the left are the target surface (a) and the source one (b) colored by electrostatic potential. The circle in (c) highlights a region where both the local geometry and the potential are similar. An area of high potential differences is near the arrow. (d) and (e) show areas of high local geometrical differences, which we render with increased transparency.*

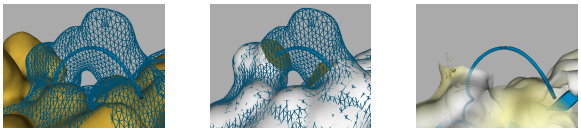


Figure 7: *Mapping of surface parts with different genus. The source mesh is pulled into the sides of the target meshes' tunnel, leading to two patches of unmapped triangles (orange parts, middle), which become transparent in the final rendering (right).*

the tunnel mesh from both sides due to the characteristics of the GVF forces. In all cases, the resulting unmapped patches represent areas which cannot be compared in a meaningful way and, hence, they are rendered transparently and omitted in the computation of the global descriptors.

7.2. Global Dissimilarity

We computed D_{sign}, D_ϕ, G, as well as the RMSD value of the aligned structures for all 152x152 pairs of conformers of the CALB ensemble (computation time about 1.5 h). Figure 8 shows a matrix plot for D_ϕ. Here, lighter patches indicate groups of conformers that exhibit similarity and can, therefore, be used for qualitative clustering. Hence, the matrix plot allows a preselection of subsets of the data set for further investigation. The plot also demonstrates that the resulting value of D_ϕ is overall symmetrical. The results for the aforementioned subset of the ensemble can be found in Figure 9. For increasing methanol activity, a change in the surface potential (both in D_ϕ and D_{sign}) can be observed. The same coherency is not visible when considering purely geometrical measurements like the RMSD and our descriptor G. The result complies with our surface renderings of the same group of data (see Figure 1).

Furthermore, we compared our global geometry descrip-

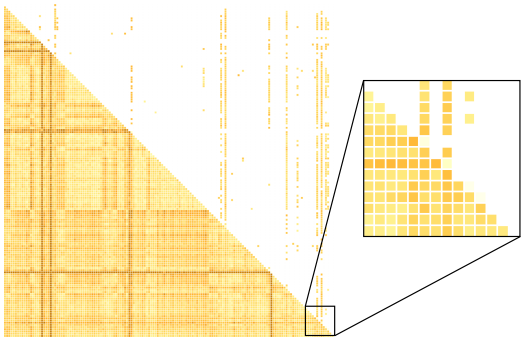


Figure 8: *A matrix plot for 152 conformers of CALB, showing the mean absolute potential difference D_ϕ. The upper half of the matrix is omitted if symmetry is given (using a tolerance value, which is defined as a percentage of to the difference range).*

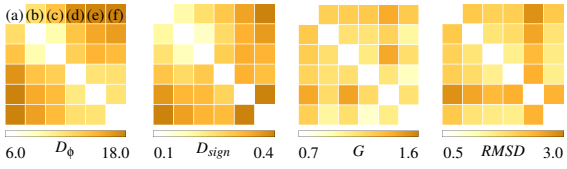


Figure 9: *2D plot of the global descriptors D_ϕ, D_{sign}, and G computed with our method, as well as the RMSD. The same subset of variants is shown as in Figure 1. The gradual change in D_ϕ and D_{sign} suggests a correlation to the changing solvent mixture that is not present in the purely geometric RMSD and G.*

tor G to the distance resulting from the 3D Zernike descriptor presented in [SLL*08] (denoted by Z) and the *RMSD* of the atomic positions. A comparison of the protein 2YPI-A with 25 other proteins was computed. The result for all three descriptors is shown in Figure 10 (with normalized scales for

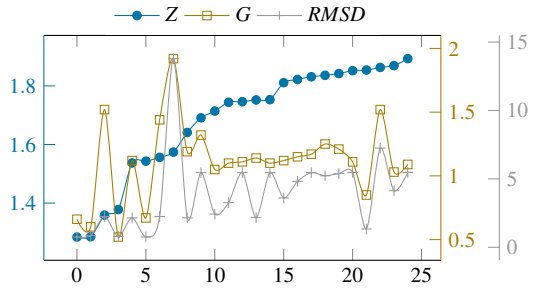


Figure 10: *Comparison of 2YPI-A with 25 other proteins using our descriptor G, the Zernike-based descriptor presented in [SLL*08] (denoted by Z), and the RMSD.*

better comparability). Our descriptor G is qualitatively different to Z, however, it resembles the *RMSD*. We assume, this is due to the fact that Z is rather sensitive to global changes, whereas our operator g integrates local deformation over the surface. The RMSD, however, is influenced by the differences in the atomic positions inside the molecule. g is insensitive to those differences since it is restricted to the molecular surface geometry. Hence, our method provides additional information compared to Z and the RMSD.

7.3. Limitations

Our mapping algorithm is intended for the detailed comparison of molecules that do not exhibit large global geometric differences. The dissimilarity measurement relies on a reasonable initial rigid alignment. For global conformation changes, our method would yield high deformation values, indicating bad comparability of the molecules. Treating these cases would require special preprocessing, e.g. a two-step deformation process, which performs global deformation prior to local deformation. With respect to molecular surface similarity, however, this is not necessary, since a comparison of molecules that exhibit large global deformations is not reasonable from a biochemists perspective.

We use an additional subdivision step to counteract the rise of oversized unmapped triangles in areas that are comparable. The subdivision, however, does not always work well with the GVF forces. In some cases, newly created vertices are immediately pulled towards the sides of very steep cavities, recreating the just subdivided overlong edges (see Figure 11). This is due to the GVF, which creates saddles in some cases [GR03]. Consequently, our algorithm in its current form cannot overcome all oversized-triangles, which can lead to asymmetrical results in the computation of G, since unmapped triangles are omitted in the integration. We, therefore, offer the user to switch between two alternative computation results, one that omits unmapped triangles, and one that incorporates them. The resulting asymmetry can be a clue to the analyst that a special case occurs here, which might be interesting for further analysis.

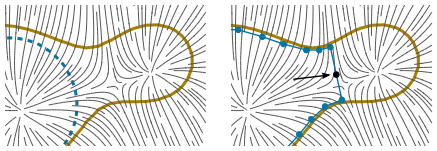


Figure 11: *In some situations, the GVF creates saddles that prevent morphing of a newly created vertex (arrow) inside the cavity.*

As an alternative to our current 2D plot, we propose to use a more elaborate matrix plot for G that shows the percentage of the surface that leads to badly matched triangles area. The height of the boxes in the matrix plot could be scaled according to the badly matched triangles area with respect to the total area of the source surface. This would allow us to complement the integrated path length for matched surface parts with the unmapped surface area.

8. Conclusion & Future Work

We presented a method for partial shape matching of surfaces using a deformable model approach. Based on a combination of rigid and non-rigid alignments, we defined a mapping that allows comparing arbitrary surface attributes. We computed local and global descriptors and visualized them in a 3D surface rendering and a matrix plot. By exemplarily applying our method to real-world biochemical data sets, we showed how it can be used for qualitative grouping and detailed analysis of molecular data.

Although our application scenario stems from the field of biochemistry, our method can be applied to different problems related to shape matching or surface attribute comparison, if appropriate rigid alignment is performed. This includes e.g. investigating properties of 2D manifolds representing vector field topology. Here, manifolds extracted from Eigenvector fields are interesting, since their properties can often not be obtained directly from the underlying vector field definition. Furthermore, our method could be applied to analyze three-dimensional medical image data (like e.g. CT scans), a field where deformable model based methods are widely used.

Acknowledgments: This work was partially funded by DFG as subprojects C.1, D.3 and D.4 of SFB 716. We thank Ł. Gricman for his input and for providing the P450 data set.

References

[BH06] BAUM D., HEGE H.-C.: A point-matching based algorithm for 3d surface alignment of drug-sized molecules. In *Computational Life Sciences II*, vol. 4216 of *Lecture Notes in Comput. Sci.* 2006, pp. 183–193. 2

[BHLW12] BOJSEN-HANSEN M., LI H., WOJTAN C.: Tracking Surfaces with Evolving Topology. *ACM Trans. Graph. 31*, 4 (2012), 53:1–53:10. 3

[Bli82] BLINN J. F.: A Generalization of Algebraic Surface Drawing. *ACM Trans. Graph. 1*, 3 (1982), 235–256. 3

[BM92] BESL P. J., MCKAY N. D.: A Method for Registration of 3-D Shapes. *IEEE Trans. Pattern Anal. Mach. Intell. 14*, 2 (1992), 239–256. 3

[BSM*06] BARGTEIL A. W., SIN F., MICHAELS J. E., GOKTEKIN T. G., O'BRIEN J. F.: A texture synthesis method for liquid animations. In *Proc. 2006 ACM SIGGRAPH/EG Symposium on Computer Animation* (2006), SCA '06, pp. 345–351. 3

[CBJ00] COSGROVE D., BAYADA D., JOHNSON A.: A novel method of aligning molecules by local surface shape similarity. *J. Comput. Aided Mol. Des. 14*, 6 (2000), 573–591. 2

[CC93] COHEN L. D., COHEN I.: Finite-Element Methods for Active Contour Models and Balloons for 2-D and 3-D Images. *IEEE Trans. Pattern Anal. Mach. Intell. 15*, 11 (Nov. 1993), 1131–1147. 4

[CK10] CAMPEN M., KOBBELT L.: Exact and robust (self-)intersections for polygonal meshes. *Computer Graphics Forum 29*, 2 (2010), 397–406. 3

[DYT05] DINH H. Q., YEZZI A., TURK G.: Texture transfer during shape transformation. *ACM Trans. Graph. 24*, 2 (Apr. 2005), 289–310. 3

[FM13] FINN P. W., MORRIS G. M.: Shape-based similarity searching in chemical databases. *Wiley Interdiscip. Rev. Comput. Mol. Sci. 3*, 3 (2013), 226–241. 2

[GP95] GRANT J. A., PICKUP B. T.: A Gaussian description of molecular shape. *J. Phys. Chem. 99*, 11 (1995), 3503–3510. 3

[GR03] GIL D., RADEVA P.: Curvature Vector Flow to Assure Convergent Deformable Models for Shape Modelling. In *In EMMCVPR* (2003), pp. 357–372. 9

[HLA04] HOFBAUER C., LOHNINGER H., ASZÓDI A.: SURFCOMP: A novel graph-based approach to molecular surface comparison. *J. Chem. Inf. Comput. Sci 44* (2004), 837–847. 2

[JM90] JOHNSON M. A., MAGGIORA G. M.: *Concepts and applications of molecular similarity*. Wiley-Interscience Publication. Wiley, 1990. 2

[Kab76] KABSCH W.: A discussion of the solution for the best rotation to relate two sets of vectors. *Acta Crystallogr. Sect. A 34*, 5 (Sept. 1976), 922–923. 3

[KS04] KRAEVOY V., SHEFFER A.: Cross-parameterization and compatible remeshing of 3D models. *ACM Trans. Graph. 23*, 3 (Aug. 2004), 861–869. 2

[KSES12] KRONE M., STONE J. E., ERTL T., SCHULTEN K.: Fast Visualization of Gaussian Density Surfaces for Molecular Dynamics and Particle System Trajectories. In *EuroVis 2012 Short Pap.* (2012), vol. 1. 5

[KWT88] KASS M., WITKIN A., TERZOPOULOS D.: Snakes: Active contour models. *Int. J. Comput. Vis. 1*, 4 (1988), 321–331. 2, 4

[LFR09] LIU Y.-S., FANG Y., RAMANI K.: Idss: deformation invariant signatures for molecular shape comparison. *BMC Bioinformatics 10*, 1 (2009), 1–14. 2

[MDA01] MONTAGNAT J., DELINGETTE H., AYACHE N.: A review of deformable surfaces: topology, geometry and deformation. *Image Vis. Comput. 19*, 14 (2001), 1023–1040. 2

[Mül09] MÜLLER M.: Fast and robust tracking of fluid surfaces. In *Proceedings of the 2009 ACM SIGGRAPH/Eurographics Symposium on Computer Animation* (2009), pp. 237–245. 3

[OF03] OSHER S., FEDKIW R.: *Level Set Methods and Dynamic Implicit Surfaces*. Applied Mathematical Sciences. Springer, 2003. 3

[PGH*04] PETTERSEN E. F., GODDARD T. D., HUANG C. C., COUCH G. S., GREENBLATT D. M., MENG E. C., FERRIN T. E.: UCSF Chimera - a visualization system for exploratory research and analysis. *J. Comput. Chem. 25*, 13 (2004), 1605–1612. 2

[PRW11] PFAFFELMOSER T., REITINGER M., WESTERMANN R.: Visualizing the Positional and Geometrical Variability of Isosurfaces in Uncertain Scalar Fields. In *Comput. Graph. Forum* (2011), vol. 30, pp. 951–960. 6

[PS09] POSTARNAKEVICH N., SINGH R.: Global-to-local representation and visualization of molecular surfaces using deformable models. In *Proc. 2009 ACM Symp. Appl. Comput.* (2009), SAC '09, ACM, pp. 782–787. 2

[Ric77] RICHARDS F. M.: Areas, Volumes, Packing, and Protein Structure. *Annu. Rev. Biophys. Bioeng. 6*, 1 (1977), 151–176. 3

[RK99] RITCHIE D. W., KEMP G. J. L.: Fast computation, rotation, and comparison of low resolution spherical harmonic molecular surfaces. *J. Comput. Chem. 20*, 4 (1999), 383–395. 2

[SAPH04] SCHREINER J., ASIRVATHAM A., PRAUN E., HOPPE H.: Inter-Surface Mapping. *ACM Trans. Graph. 23*, 3 (Aug. 2004), 870–877. 2

[SB08] SCHNEIDER G., BARINGHAUS K.-H.: *Molecular design: concepts and applications*. Wiley-VCH, 2008. 3

[SCP12] SOTIRAS A., CHRISTOS D., PARAGIOS N.: *Deformable Medical Image Registration: A Survey*. Research Report RR-7919, INRIA, Sept. 2012. 2

[SDP13] SOTIRAS A., DAVATZIKOS C., PARAGIOS N.: Deformable Medical Image Registration: A Survey. *IEEE Trans. Med. Imaging 32*, 7 (2013), 1153–1190. 2

[SHL*11] SHEN T., HUANG X., LI H., KIM E., ZHANG S., HUANG J.: A 3D Laplacian-driven parametric deformable model. In *ICCV* (2011), IEEE, pp. 279–286. 5

[SLL*08] SAEL L., LA D., LI B., RUSTAMOV R., KIHARA D.: Rapid comparison of properties on protein surface. *Proteins Struct. Funct. Bioinforma. 73*, 1 (2008), 1–10. 2, 8, 9

[TMA*06] TANG J., MILLINGTON S., ACTON S. T., CRANDALL J., HURWITZ S.: Surface extraction and thickness measurement of the articular cartilage from mr images using directional gradient vector flow snakes. *IEEE Trans. Biomed. Engineering 53*, 5 (2006), 896–907. 3

[TV08] TANGELDER J. W., VELTKAMP R. C.: A survey of content based 3D shape retrieval methods. *Multimed. Tools Appl. 39*, 3 (Sept. 2008), 441–471. 2

[TWK88] TERZOPOULOS D., WITKIN A., KASS M.: Constraints on deformable models: recovering 3D shape and nongrid motion. *Artif. Intell. 36*, 1 (Aug. 1988), 91–123. 2, 4

[VFBHC00] VIA A., FERRÈ F., BRANNETTI B., HELMER-CITTERICH M.: Protein surface similarities: a survey of methods to describe and compare protein surfaces. *Cell. Mol. Life Sci. C. 57*, 13-14 (2000), 1970–1977. 2

[vKZHCO11] VAN KAICK O., ZHANG H., HAMARNEH G., COHEN-OR D.: A Survey on Shape Correspondence. *Comput. Graph. Forum 30*, 6 (2011), 1681–1707. 2

[WCX*13] WENG Y., CHAI M., XU W., TONG Y., ZHOU K.: As-rigid-as-possible distance field metamorphosis. *Computer Graphics Forum 32*, 7 (2013), 381–389. 3

[WTGT10] WOJTAN C., THÜREY N., GROSS M., TURK G.: Physics-inspired topology changes for thin fluid features. *ACM Trans. Graph. 29*, 4 (July 2010), 50:1–50:8. 3

[XP98] XU C., PRINCE J. L.: Snakes, shapes, and gradient vector flow. *Image Process. IEEE Trans. 7*, 3 (1998), 359–369. 3, 4, 5