

Real-Time Molecular Visualization Supporting Diffuse Interreflections and Ambient Occlusion

Robin Skånberg, Pere-Pau Vázquez, Victor Guallar and Timo Ropinski *Member, IEEE*

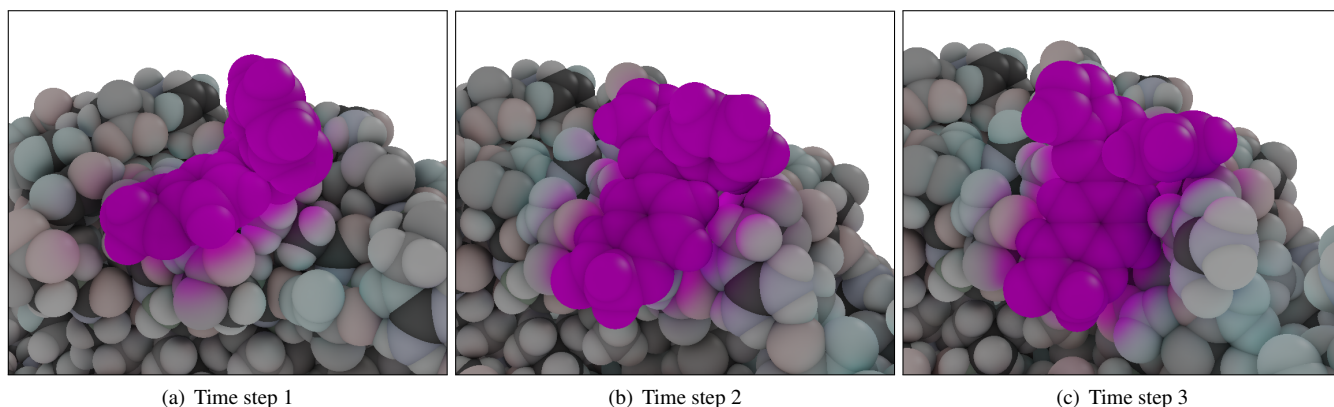


Fig. 1. By deriving analytic expressions, we can enhance molecular visualizations and realize interreflections in real-time. The images (a-c) show three time steps of a molecular simulation investigating the interaction between a magenta-colored ligand and a desaturated receptor molecule, which receives exaggerated diffuse interreflections. Due to these interreflections it can be seen how the ligand enters the active site.

Abstract—Today molecular simulations produce complex data sets capturing the interactions of molecules in detail. Due to the complexity of this time-varying data, advanced visualization techniques are required to support its visual analysis. Current molecular visualization techniques utilize ambient occlusion as a global illumination approximation to improve spatial comprehension. Besides these shadow-like effects, interreflections are also known to improve the spatial comprehension of complex geometric structures. Unfortunately, the inherent computational complexity of interreflections would forbid interactive exploration, which is mandatory in many scenarios dealing with static and time-varying data. In this paper, we introduce a novel analytic approach for capturing interreflections of molecular structures in real-time. By exploiting the knowledge of the underlying space filling representations, we are able to reduce the required parameters and can thus apply symbolic regression to obtain an analytic expression for interreflections. We show how to obtain the data required for the symbolic regression analysis, and how to exploit our analytic solution to enhance interactive molecular visualizations.

Index Terms—Molecular visualization, diffuse interreflections, ambient occlusion

1 INTRODUCTION

Modern molecular simulation algorithms, such as molecular dynamics or Monte Carlo solvers, enable scientists to capture the interaction of complex molecules. Such state-of-the-art simulations are for instance used to investigate the binding nature of new pharmaceuticals or the turnover rate of genetically modified enzymes. For all these protein-ligand interactions it is crucial to analyze how the two substances form the resulting complex, to understand their interplay and to be able to support an affinity-driven ligand modification.

Unfortunately, the visualization of these substance compounds is challenged by their structural complexity. As a consequence, despite that their exact interaction needs to be studied, the spatial arrangement can often not be inferred from current state-of-the-art visualizations.

Within this paper we propose a novel approach for visualizing molecular structures whereby we focus on the atom level. The presented method has been developed with the motivation to better communicate spatial arrangements as well as non-spatial simulation parameters, which play an important role when observing molecule ligand interactions. While state-of-the-art molecular visualization techniques simulate ambient occlusion (AO) to better communicate structures, we aim at capturing the interreflections between molecule and ligand to communicate their spatial arrangement. Furthermore, we exploit modifications of these interreflections to visualize pairwise atom interaction strengths, which give important insights for affinity-driven ligand modification. To enable these visualizations, we capture mutual illumination, resulting from the reflection of light from one object onto another object.

Especially in cavities or enclosures, as they also occur in molecular structures, mutual illumination can provide a significant component of surface illumination [28]. It is not only known that such interreflections play an important role in the perception process [2], but also known that they resolve the classic bas-relief ambiguity and thus have a direct influence on shape perception [5]. Within this paper, we apply mutual reflections to molecular structures, and we further show how these interreflections can be modified to interactively visualize atom interaction strengths, which tell the domain expert which atom interactions between molecule and ligand are most crucial for the binding process. Thus, the presented techniques have been motivated with the aim to help the researchers in understanding two important factors that play a role in the simulation: the spatial arrangement of the molecules, and the forces that interact in each configuration.

- Robin Skånberg is with the Visual Computing Group, Ulm University. E-mail: robin.skanberg@uni-ulm.de.
- Pere-Pau Vázquez is with the MOVING Group, UPC Barcelona. E-mail: ppau@cs.upc.edu.
- Victor Guallar is with the Barcelona Supercomputing Center. E-mail: victor.guallar@bsc.es.
- Timo Ropinski is with the Visual Computing Group, Ulm University. E-mail: timo.ropinski@uni-ulm.de.

Manuscript received 31 Mar. 2015; accepted 1 Aug. 2015; date of publication 20 Aug. 2015; date of current version 25 Oct. 2015.
For information on obtaining reprints of this article, please send e-mail to: tvcg@computer.org.
Digital Object Identifier no. 10.1109/TVCG.2015.2467293

To achieve our goals, we present a novel approach for visualizing molecular structures, which not only captures AO, but also diffuse interreflections. While approximations for interactive AO effects exist for molecular visualization, e.g., [17], there are no techniques that support interreflections. To realize high quality mutual interreflections, we exploit an analytical approach that is based on symbolic regression analysis of the light interactions occurring in space filling representations of molecules [44], such as van der Waals (vdW), solvent-accessible surfaces (SAS), and solvent-excluded surfaces (SES). To analyze these light interactions, we exploit path tracing as a physically correct light transport solution in molecular structures. By sweeping carefully determined visualization parameters, we can use our path tracer to generate interreflection data, based on which we are able to derive multivariate analytical expressions. These expressions can then be evaluated in real-time and thus enable complex illumination effects in molecular visualization at interactive frame rates without requiring additional memory (see Figure 1). Thus, the time-varying results of molecular simulations can be visualized, while exploiting interreflections to improve spatial comprehension as well as encoding important simulation parameters. To achieve this, we make the following contributions in this paper:

- We present a novel visualization algorithm for space filling representations of molecular structures, which analytically realizes ambient occlusion and diffuse interreflections.
- We propose how to use mutual interreflections as a new visual communication channel for visualizing pairwise interaction strengths, to guide domain experts to affinity-driven ligand modifications.
- We introduce symbolic regression into the visualization field as a means to analytically capture multidimensional functions to enable interactive visualization.

The paper is structured as follows. After discussing related work in Section 2, we outline our approach for analytically representing interreflections on an atomic scale in Section 3. In Section 4 we discuss how to acquire the multi-parametric data required for our data-centric approach, and how we exploit symbolic regression to obtain analytic expressions, that can be evaluated in real-time. The actual visualization approach is described in Section 5, while results are discussed, both from a quality and a performance point of view, in Section 6. The paper concludes in Section 7, where we also discuss future work.

2 RELATED WORK

Besides the challenges related to communicating complex molecular structures [6, 14, 16, 19, 21, 47], visualization algorithms need to be able to show the composition of elements and the various reactions that happen in simulations. In the following, we classify the methods described in literature into two groups: methods tailored to comprehend atomic structures and methods for illustrating molecular reactions. After discussing these we will focus on related work in simulating global illumination.

Atomic structure visualization. Many authors focused on massive atomic model rendering [24, 39], such as LOD approaches [11, 30, 37], instancing [12], or billboards and glyphs [13, 17, 45, 40]. Grottel et al. described an approach which exploits two-level coherent culling to render large-scale molecular dynamics results [18]. Coarse culling exploits hardware occlusion queries, while maximum depth mipmaps are used to achieve per-vertex culling. More recently, Lindow et al. described a method that enables interactive visualization of atomic data at different scales [32]. They exploit the fact, that atomic data is often rendered opaquely and that it often contains reoccurring structures, which lets them exploit instancing. Zwan et al. concentrate on the continuous transition between different visualization motifs for molecules, e.g. from space filling to ribbon rendering [46]. Lawonn et al. add feature lines, and hatching to emphasize features like cavities, channels, and pockets [29]. Cipriano et al. also add hatching effects by analyzing surface properties [8]. Another popular method to facilitate spatial comprehension is the inclusion of AO [17, 27, 45]. Krone et al. perform depth darkening [34] when real-time ray casting molecules using a surface-based model [27]. Grottel et al. improve on the previous method by computing AO in object space for particle

models [17]. In order to do so, a grid is computed that stores the occupancy information derived from the neighboring atoms, which enables AO in real-time. Kottraval et al. introduces depth-of-field techniques to molecular rendering, which provides additional depth cues to the user [25]. To speed up rendering, many authors exploit deferred shading techniques in molecular visualization [13, 18, 32]. As a thorough review of all atomic structure visualization techniques is beyond the scope of this paper, we would like to refer to the state-of-the-art report on molecular structure visualization by Kozlíková et al. [26].

Illustrating molecular reactions. Falk et al. illustrate the different transport modes in mitogen-activated protein kinase signaling molecules (MAPK) within a cell [12]. Molecular trajectories and reactions are identified by augmenting geometric objects, such as paths that illustrate molecule trajectories, and arrows that depict reactions. Sarikaya et al. visualize the performance of protein classifiers by applying different colors to the protein chains on which the classifiers have operated [42]. Their inspection mode consists of two views, one containing the molecule divided in small multiples, and a second view where the surface of the molecule is color encoded with the information on the predictions of the classifier. They also provide illustrative cues to enhance the perception of atoms. More concretely they add contour lines and tooned AO. Khazanov and Carlson provide several visualization elements in the form of tables and tabular graphs [23]. They also visualize the contacts with ligands and their binding sites by modifying the atoms' vdW radii and color proportionally. This depiction is performed after the whole simulation and individually for each amino acid. More recently, Scharnowski et al. have compared the differences between surface features of the molecules, e.g., electrostatic potential, using deformable models [43]. Grottel et al. visualize the values of electrostatic dipoles by using a color map overlapping the molecular structures [15]. Cipriano and Gleicher [7] illustrate the charges over the molecular surface, but they do so by stylizing both the surface shape and the charge values to facilitate comprehension. Some software packages provide a means to overlap a set of semi-transparent spheres around atoms, used to color encode atom properties, e.g., Coulomb charges or hydrophobicity, such as in VESTA [35, 36]. However, today no visualization algorithm communicates the forces acting between all atoms in real-time, without introducing visual clutter. Finally, other visualization methods illustrate path channels [31] or molecular motion by the use of glyphs [3].

Simulating Global Illumination. As our technique simulates global illumination, we briefly discuss related algorithms in this field. Ramamoorthi et al. describe an efficient technique for storing and retrieving low-frequency irradiance maps using spherical harmonics [38]. Further extensions to this technique exist but they all rely on precomputing the irradiance of the environment. This limits the use cases to scenes where the irradiance can be precomputed. Ren et al. describes an approach that uses radiance regression functions to compute global illumination using a neural network [41]. The underlying idea is similar to our approach but the method requires extensive pre-processing per scene and is limited to static geometry. Bunnell presents a technique for simulating ambient occlusion and indirect light between triangles by using oriented discs as an underlying representation [4].

3 ANALYTIC ATOM-SCALE INTERREFLECTIONS

Determining an analytical representation of mutual reflections on an atomic scale is a challenging task as these light interactions follow the complexity of the recursive rendering equation [22]. It means that, for each scene element, we have to take into account all other scene elements, which contribute to the visual appearance through interreflections. The recursive nature of the rendering equation leads to an even increased complexity, as color bleeding effects can result from several bounces of light between the involved structures. Capturing this complexity in real-time is challenging, and does not scale well with the number of atoms comprising a complex molecule. Therefore, we exploit an analytical solution to the rendering equation as applied to space filling representations, e.g., vdW, SAS, SES, of molecular structures. To be able to obtain such an analytic representation, we take into account the following observations:

- Interreflections can be computed on an atomic scale rather than per surface point (*Observation 1*).
- Due to the spherical shape of atoms and the nature of the rendering equation, reflection changes on a surface of an atom can be

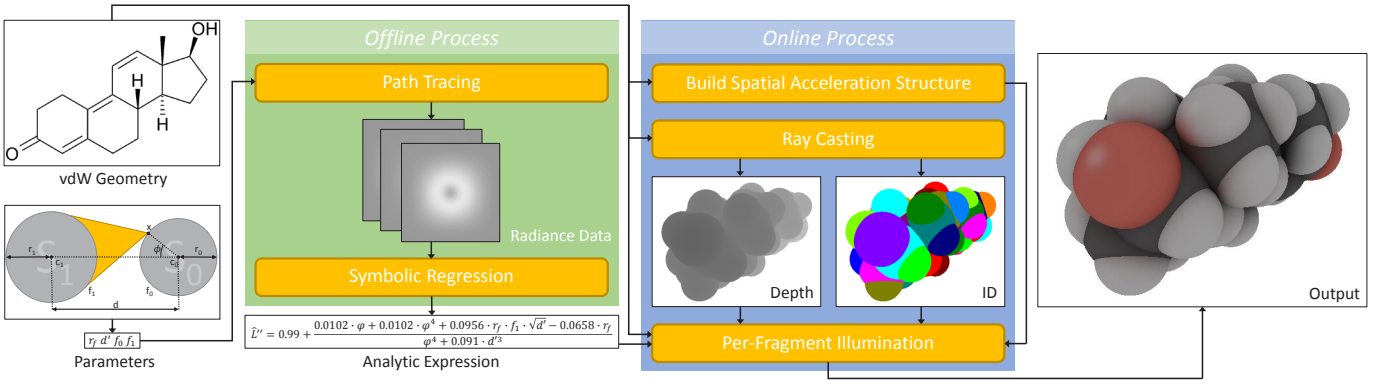


Fig. 2. An overview of our algorithm and the underlying workflow: The green box represents a pre-processing step, usually done only once, where an expression is fit to path traced data. The blue box shows the process which occurs per frame in order to render a vdW geometric representation of a molecule on screen.

considered low-frequency illumination effects (*Observation 2*).

- Reflections resulting from several adjacent structures can be reconstructed by modulating the reflection contributions of pairwise combinations of these structures (*Observation 3*).

Observation 1 holds, as in the targeted space filling representations, atoms are inherently represented as spheres. This can be even assumed for the SES representation, as it can be extracted using the rolling ball metaphor. However, in this case it needs to be taken into account that, while some of the patches resemble a sphere, others are concave. When considering the light interactions between two surface points on the spheres representing two atoms, it is clear that the points' location, and thus their light interaction, can be directly derived from the properties of the atom, i.e., the location and the vdW radius. As a consequence *Observation 1* holds, and we can compute interreflections on an atom level. *Observation 2* directly follows from how light interactions are described by the rendering equation. Due to the continuous integration and the weighting with the continuous cosine term, mutual interreflections can be considered low frequency illumination effects. Thus, we can make *Observation 2* and see it as an indicator that an analytic expression exists, which captures such reflections. Finally, *Observation 3* holds due to the additive nature of light, which tells us that we can treat different illumination sources independently. Thus, *Observation 3* enables us to analyze pairwise atomic light interactions to determine the desired analytic expression. It shall be noted, that *Observation 3* results in an approximation only, as it does not consider light bouncing off several times at other objects, before coming back to the initial object. However, in perceptual experiments it has been shown that similar effects resulting from approximate visibility can be neglected [48].

Algorithm Overview. Due to the three observations discussed above, we are able to find an analytic function which captures diffuse interreflections in molecular scenes. Therefore, we exploit a generalized molecular scene, which consists of two atoms only. These two atoms have varying radii, reflectances and distances in between, such that the scene can be fully defined by a few parameters only. To find the desired analytical function, we path trace different instances of this generalized scene, which only vary with respect to these parameters. As a result we receive a set of radiance maps which captures the diffuse interreflections for the path traced scene instances. This radiance data is then used as input to a machine learner to find the analytic function, which depends on the parameters specifying the generalized scene. During rendering, we can then take into account atom pairs to derive these parameters for the current molecule, before feeding these parameters into the analytic function in order to simulate diffuse interreflections.

A high-level overview of this process is given in Figure 2. First as an offline pre-processing step, radiance maps are generated from a path tracer using the scene parameters as input. These maps are then used by the symbolic regression machine learner which fits an expression to the data. Note that this offline step only has to be done once, as it produces a reusable expression that captures the diffuse interaction within the given parametric space. In the second online process, which is done per frame, space filling geometry is placed in a spatial acceleration structure and rendered in a first geometric pass where the ray-casted depth as

well as the atom ID is stored. In the second render pass, the actual illumination is computed per fragment using the spatial acceleration structure to find the neighboring spheres and applying the expression.

4 DATA GENERATION AND SYMBOLIC REGRESSION

Within the next subsection, we describe how we exploit a specially parametrized path tracer to solve the rendering equation for our generalized molecular scene, and thus capture the illumination data describing the interreflections. In Section 4.2, we then describe how we exploit symbolic regression to find an analytical description of the interreflection behavior.

4.1 Capturing Illumination Data

Before we are able to analytically describe the interreflection between atoms, we need to capture these interreflections in a physically-based manner. The rendering equation has been proposed as a physically-inspired equation by Kajiya to describe light transport in a geometric scene [22]. It is based on the laws underlying heat transfer, and can thus be used to derive the radiance L that leaves a single surface point x in the direction of the viewer $\vec{\omega}_o$ in dependence of the geometric structures contained in the scene:

$$L(x, \vec{\omega}_o) = L_e(x, \vec{\omega}_o) + \int_{\Omega} f(x, \vec{\omega}_i, \vec{\omega}_o) \cdot \cos(\theta) \cdot L_i(x, \vec{\omega}_i) d\vec{\omega}_i. \quad (1)$$

Here, $L_e(x, \vec{\omega}_o)$ describes the radiance emitted from x into direction $\vec{\omega}_o$, and $f(x, \vec{\omega}_i, \vec{\omega}_o)$ stands for the BRDF relating the radiance reaching x from $\vec{\omega}_i$ to the outgoing direction $\vec{\omega}_o$. The multiplication with $\cos(\theta)$ results in energy conservation, as it weights the incoming radiance, determined by recursive evaluations denoted as L_i , for light reaching x from all incoming directions $\vec{\omega}_i$.

Atom interreflections. To capture the interreflections between atoms, we use a slightly modified version of Equation 1. As we aim at only capturing interreflections between atoms, and not their light contributions, we eliminate the emissive contributions L_e from Equation 1, and thus only solve the integral over $\vec{\omega}_i$:

$$L(x, \vec{\omega}_o) = \int_{\Omega} f(x, \vec{\omega}_i, \vec{\omega}_o) \cdot \cos(\theta) \cdot L_i(x, \vec{\omega}_i) d\vec{\omega}_i. \quad (2)$$

However, Equation 2 would not have any energy contributions, since L_i is just a recursive evaluation of L . Therefore, we introduce environmental lighting as the only source of energy in the scene. When observing Equation 2, we can see that the radiance leaving x in direction $\vec{\omega}_o$ is dependent on the location and the reflective properties of x , as well as all other points in the scene. Thus, if we could determine a closed form expression $\hat{L}(x, \vec{\omega}_o, \{x_0, \dots, x_i\}, \{f_0, \dots, f_i\})$, that describes the radiance at x dependent on the scene geometry and the reflectance functions, such an expression could be directly used for rendering. Obviously, finding such an expression is impractical for arbitrary complex scenes. Fortunately, when dealing with space filling representations of molecules the situation is much easier. When dealing with vdW,

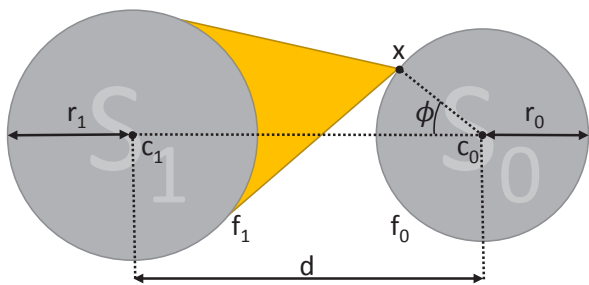


Fig. 3. The parametrization used to solve $\hat{L}(\cos(\phi), \vec{\omega}_o, d, r_f, f_0, f_1)$ with a path tracer. The yellow triangle shows the area influencing x when only a single bounce would be considered.

SAS, or SES, the three space filling representations most widely used in molecular visualization, we can understand the scene as a set of similar building blocks. For vdW and SAS these building blocks are directly given as spheres, a fact that can be exploited to simplify finding a closed form expression as expressed by *Observation 1* (cf. Section 3). Thus, within this subsection, we discuss how to do this for vdW and SAS, while we present how to extend this concept to SES in Section 5.4.

Due to *Observation 3*, we can initially limit ourselves to pairwise interreflections, and thus consider two atoms only. As a consequence, the closed form expression \hat{L} can be simplified to be dependent on fewer parameters. When assuming that the reflective properties for each individual atom are uniform, \hat{L} would depend only on two reflection functions f_0 and f_1 instead of a larger set. Furthermore, the set of all points describing the scene geometry could be simplified to those describing the surfaces of the two atoms. Nevertheless, when considering a sufficiently high resolution, such a parametrization would still be impractical. Therefore, we exploit *Observation 1*, which enables us to simplify the set of all points to the center c and the radius r of the spheres representing an atom. The closed form expression \hat{L} would now be dependent on the following parameters $\hat{L}(x, \vec{\omega}_o, c_0, c_1, r_0, r_1, f_0, f_1)$, which are illustrated in Figure 3. As we are considering two isolated atoms, we can further simplify \hat{L} , by only considering the distance d in between c_0 and c_1 . Furthermore, it is sufficient to take into account the ratio of the radii $r_f = \frac{r_0}{r_1}$. Accordingly, the number of parameters for \hat{L} can be further simplified such that we obtain $\hat{L}(x, \vec{\omega}_o, d, r_f, f_0, f_1)$. Finally, we are able to exploit the radial symmetry of the spheres, such that we only need to compute the radiance distribution along one Longitude, which is parametrized by the polar angle ϕ . We can use this parameter to replace the actual position x with the cosine of the angle ϕ between $c_0 - x$ and $c_0 - c_1$ (see Figure 3), and thus obtain $\hat{L}(\cos(\phi), \vec{\omega}_o, d, r_f, f_0, f_1)$. To obtain a single view-independent value, we integrate the radiance over all directions $\vec{\omega}_o$ in the hemisphere Ω and obtain:

$$\hat{L}'(\cos(\phi), d, r_f, f_0, f_1) = \int_{\Omega} \hat{L}(\cos(\phi), \vec{\omega}_o, d, r_f, f_0, f_1) d\vec{\omega}_o. \quad (3)$$

\hat{L}' is now only dependent on five parameters. All these five parameters are directly accessible during rendering, as they can be easily computed for each pair of atoms. Here, d and r_f are global geometric atom parameters, which are computed by taking into account both atoms. f_0 and f_1 are global parameters which can be obtained from a single atom, and $\cos(\phi)$ is a local parameter which needs to be determined for each fragment position x to be shaded.

Interreflection parameters. As we aim at finding an analytic expression capturing \hat{L}' , we sample \hat{L}' for different parameter sets, to initially get a data-based representation. To realize such a parameter sweep, we employ path tracing which has been proposed as the first solution to the rendering equation [22]. For the data generation, we use a scene made up of two spheres with the centers c_0 and c_1 , the radii r_0 and r_1 as well as the reflective properties f_0 and f_1 (cf. Figure 3). However, instead of projecting onto the image plane, we have modified the path tracer such that it computes the radiance distribution on the surface of the atom represented by sphere S_0 . Furthermore, we

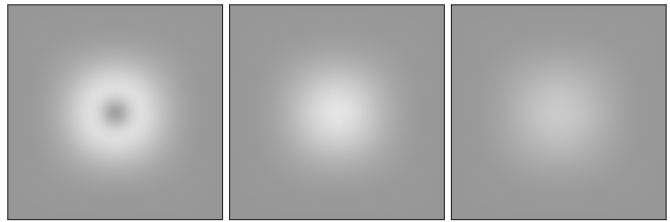


Fig. 4. Examples of spherical radiance maps as generated by our path tracer as the input to the symbolic regression analysis. The examples vary horizontally from left to right with respect to $d' \in [1.05, 1.30, 1.55]$, with $r_f = 1.0$ and $f_0 = f_1 = 0.8$. The radiance maps are cropped to $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$.

compute radiance in an achromatic manner for one channel only, to be able to apply it flexibly to different colored materials during interactive visualization. Figure 4 shows three examples of such spherical radiance maps as written out by our path tracer. The maps show the radiance on the sphere representing atom a_0 , whereas ϕ varies vertically, and θ varies horizontally. The theta values are cropped to the range $[-\frac{\pi}{2}, \frac{\pi}{2}]$. The shown spherical radiance maps vary horizontally from left to right with respect to $d' \in [1.05, 1.30, 1.55]$, whereby $d' = \frac{\|c_0 - c_1\|}{r_0 + r_1}$. We have chosen d' as a unit-less ratio, as it enables us to use the same derived expression also for problems at other scales. The effect of distance can be observed on the spherical radiance maps. While the map for $d' = 1.05$ clearly show occlusion effects in the center, with an increase in distance the interreflections resulting from multiple bounces take over. More details become visible when analyzing the radiance for the point x closest to S_1 and lying on S_0 . To analyze the effects of AO and diffuse interreflections independently, we have path traced a scene consisting of two equally sized spheres with a varying distance d . The plots in Figure 5 illustrate this behavior for six different values for d with $d \in [2.05, 2.15, 2.25, 2.35, 2.45, 2.55]$, whereby the degree of redness of a curve is inversely proportional to the distance d between S_0 and S_1 . The three plots show the radiance, normalized over the uniformly lit back side of S_0 , when using a regular path tracer (Figure 5 (a)), path tracing when only considering AO but no multi-bounce interreflections (Figure 5 (b)), and path tracing with multi-bounce interreflections where the AO contribution has been eliminated (Figure 5 (c)). As it can be seen in Figure 5 (a), under a certain distance the AO contribution overrides the diffuse interreflections. This effect results from the radiance decrease through AO as illustrated in Figure 5 (b). When looking at Figure 5 (c), where we have eliminated the AO contribution, it becomes clear that not only AO must be considered when simulating light interactions, but also that diffuse interreflections have a significant impact on the visual result.

Since spherical radiance maps inherently store the radiance in dependence of ϕ , we do not have to take this parameter into account when performing our parameter sweep. For the remaining four parameters, we have to first determine the parameter space, before we can sample it uniformly. To determine realistic bounds for the ratio of radii r_f , we have taken into account the calculated atomic radii as they may arise in the molecules to be visualized. Clementi et al. determined

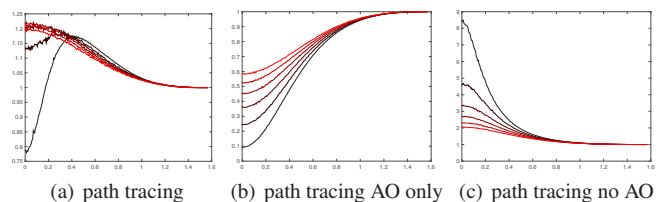


Fig. 5. Radiance obtained from a path tracer (a), a path tracer with AO only (b), and a path tracer with eliminated AO (c). The radiance is shown in relation to the distance d of two atoms. The plots show the normalized radiance for $d \in [2.05, 2.15, 2.25, 2.35, 2.45, 2.55]$, whereby the degree of redness of a curve is inversely proportional to d .

based on a theoretical model that the atomic radii range from $120pm$ for hydrogen to $298pm$ for caesium [9]. When discarding the less relevant atoms having larger radii, we can derive the following value range for $r_f \in [0.4, 2.5]$. Since the value range for the distance between

the two atoms has been chosen to be a unit-less as $d' = \frac{\|c_0 - c_1\|}{r_0 + r_1}$, we can eliminate a scale in this measure, and are able to apply it invariant of the scale of the underlying scene. The effective range for d' has then be chosen as $d' \in [0.5, 2.0]$, providing a realistic distance range when considering light falloff. A value of $d' < 1.0$ results in the atoms intersecting, which leads to some pixels being completely black in the output image. These pixels are masked out as they impede the model-fitting through the symbolic regression by introducing unwanted data and high frequencies, which should not be present (*Observation 2*). For the reflectance values f the parameter range equals $[0.1, 0.9]$. We have chosen 0.1 as lower bound, as perfectly black atoms would not result in any interreflections. The upper bound of 0.9 is motivated by the fact, that we would like to avoid perfectly reflecting materials, which would decrease the signal-to-noise-ratio in the path tracer results. Based on these derived parameter ranges, all four parameters can be swept, based on a uniform sampling of the parameter space given by r_f , d' , f_0 , and f_1 . We have chosen to generate $19 \times 12 \times 9 \times 9 = 18468$ spherical radiance maps, with the uniform sampling. Due to the modest variance of the data in relation to the computation time required to generate the data, we found this size to be adequate. Thus, we obtained a multi-parametric data set which describes the interreflections of two atoms dependent on r_0 , r_1 , c_0 , c_1 , f_0 and f_1 .

4.2 Symbolic Regression

After we have generated the multi-parametric spherical radiance map data set as described in the previous section, we need to obtain an analytic expression for \hat{L}' to be able to realize the desired interreflections in real-time. While in principle many different machine learning techniques could be used to find such an analytic function, we have decided to exploit symbolic regression [44]. Symbolic regression has two main benefits, which make it the optimal choice for the task at hand. First, as it is an unbiased approach, we do not need to input any previous knowledge about the analytic function to be found. Thus, we do not need to formulate hypothesis on how the result behaves, which at the same time could influence the result in an unexpected behavior. The second benefit of symbolic regression lies in the fact, that it is realized through genetic programming. By iteratively applying mutations to generate analytic expressions with a better fit of the data, symbolic regression does not only output the expression with the best fit, in our case the absolute error normalized over all expressions, but also the intermediate results. Finally, symbolic regression can be expected to generate more accurate and simpler expressions than traditional machine learning techniques, such as neural networks or support vector machines [44]. In the context of this paper, we have used the Eureqa software package, which implements symbolic regression as described by Schmidt and Lipson [44]. Figure 6 shows a plot depicting the output of the symbolic regression run identifying the analytic expressions used for generating the visual results in this paper. It can be seen, that the initially fit expression results in an absolute error of 0.36, and that it took 21 iterations, the number of solutions depicted in the plot, to obtain an expression with an absolute error as low as 0.14. It can also be seen that while the error is minimized, the complexity of the resulting expression increases from 1 to 44. The complexity is calculated based on the absolute number of used variables and mathematical operators, as well as the relative complexity of the mathematical operators. Since not only the accuracy, but also the complexity of each result can be retrieved, multiple analytic expressions can be used to realize LOD effects in the rendering stage of the visualization pipeline, and the iterative output of results having different levels of complexity is a perfect fit with LOD techniques applied in molecular visualization [37].

Providing a detailed explanation of the symbolic regression algorithm would be beyond the scope of this paper. Therefore, we limit ourselves to a brief outline of the basic principles, and refer to the original work by Schmidt and Lipson for a more thorough discussion [44]. The algorithm works by initially generating random candidate symbolic functions, i.e., analytic expressions describing the underlying principles of the data to be fit. During the candidate exploration phase, the algorithm derives symbolic partial derivatives of pairs of parameters, i.e., in our case r_f , d' , f_0 , and f_1 , for each candidate expression. Afterwards,

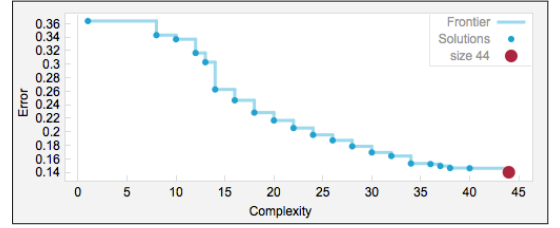


Fig. 6. Plot showing the accuracy, i.e., the absolute error, of the solutions retrieved from the symbolic regression analysis versus the complexity. In total 21 solutions have been generated in a time frame of roughly 9 hours.

the predicted partial derivatives are compared to the numerical partial derivatives, in order to select the best fitting equations for the next iteration. These are then changed by applying small variations, and serve as input for the next iteration. Thus, iteratively analytic expressions are generated which have a better fit, but also become more complicated to evaluate.

When applying this strategy to our multi-parametric spherical radiance map data set, 21 iterations led to the following analytic expression capturing \hat{L}' with a fit of 0.385, i.e., the normalized absolute error, at a complexity level of 44:

$$\hat{L}_0'' = 0.99 + \frac{0.0102 \cdot \phi + 0.0102 \cdot \phi^4 + 0.0956 \cdot r_f \cdot f_1 \cdot \sqrt{d'} - 0.0658 \cdot r_f}{\phi^4 + 0.091 \cdot d'^3} \quad (4)$$

We have obtained Equation 4 after roughly 9 hours of computation on an Intel Core i5, 3.2GHz Quad-Core System with 8 GB of RAM. As it was the result with the best fit, with a manageable complexity, we have used this function to generate the visual results throughout this paper. An interesting observation about Equation 4 is the fact, that it does not depend on f_0 , the parameter describing the reflectivity of the current atom a_0 . This again underlines the importance of the unbiased approach realized through symbolic regression, since we have initially deemed this parameter as relevant. Otherwise, we would not have taken it into account in the path tracing process. While Equation 4 presents with 0.385 a reasonably good fit to the data, at the same time it results in a smoothed out version of \hat{L}' . This is beneficial as it helps to eliminate the noise typically present in path traced images.

5 INTERACTIVE VISUALIZATION

Now that we have derived analytic expressions capturing atomic interreflections, we can evaluate the expressions during visualization. Therefore, for each atom we need to know during rendering which atoms are in its vicinity. Based on these atoms, we can derive the parameters necessary to evaluate \hat{L}_0'' , i.e., r_f , d' , f_0 , and f_1 . We first describe how to extract the atoms in the vicinity of the current atom in Section 5.1. Once this information is obtained, the light interactions between these and the current atom can be evaluated. To be able to use diffuse interreflections as a separate visual communication channel, we have decided to additionally add scalable AO during rendering time. Thus, we first describe how to exploit the results of symbolic regression to compute diffuse interreflections in real-time (see Section 5.2), before we discuss how to independently emphasize AO effects (see Section 5.3). Finally, we present how the developed techniques can be applied to SES (see Section 5.4), which requires a few modifications as compared to vdW and SAS.

5.1 Fixed-Radius Neighbors Search

Determining the neighbors of an atom is crucial, in order to be able to derive the distance d' as well as the ratio of radii r_f . In fact the problem to be solved is a variant of the radius neighbor search problem [1], which finds the set of all neighbors within a given radius:

$$N(a_0, r) = \{a_i \in A \mid \|c_0 - c_i\| < r\}, \quad (5)$$

where a_0 is the current atom with the center at c_0 , and A is the set containing all atoms a_i with their center at c_i . $r \in \mathbb{R}$ is an arbitrary radius around a_0 . Since r is fixed in our case the problem to be solved

can be simplified to the fixed-radius neighbors search problem. This variant is easier to be solved, as the knowledge about r can be taken into account when building up a query data structure. It could be shown, that grid-based techniques are a good solution to this problem, because they support localization of points in constant time, while minimizing the overhead when inspecting neighboring cells [1].

As performance is critical, we have decided to adopt one of the most recent grid-based techniques for fixed-radius neighbor search, which has been proposed by Hoetzlein [20]. Applied to our use case, the technique can be best described as follows. First, the world space containing all atoms is uniformly divided into grid cells. After this division, each atom a_i is placed in the respective grid cell determined by its center c_i . Based on this placement step, each atom a_i gets a global index id_g and a local index id_l assigned to it. While id_g specifies the index of the cell the atom has been placed in, id_l is dependent on the order the a_i have been processed, and tells how many atoms have been assigned before to the same cell. Thus, after this first placement step, we obtain an unordered atom list, whereby id_g and id_l are assigned to each atom. In a second step, the retrieved atom list is sorted based on id_g , such that the id_l atoms belonging to each cell are stored subsequently. Based on this list and the id's stored with the atoms, we are able to fetch the neighboring cells in constant time, while it takes linear time with respect to the number of atoms per cell to obtain all neighbors stored in a cell.

The technique has originally been developed for particle simulations, but since the number of atoms is rather low compared to the number of particles in state-of-the-art particle simulations, this technique provides a very efficient solution. A critical aspect in our application scenario is to choose an adequate cell size. As we know what area of influence shall affect the interreflections on an atom, we should choose the cell size such that all atoms contributing to the interreflection on a_i are contained in a cell adjacent to the one containing a_i . Thus, we can limit the neighboring cell access to 27 cells (3^3) – 26 adjacent and the center cell. Otherwise, 125 cells would need to be inspected, which would result in a data access bottleneck. Within this paper, we use three different cell sizes. To capture diffuse interreflections, which are very local effects, we have chosen a cell size of 2\AA or 6\AA , which resulted in pleasing visual results. Instead, to convey the interaction strength for atoms, we have chosen a cell size of 12\AA . As molecular dynamics simulations traditionally use 12\AA as a neighborhood region, we have decided to determine the grid cell size to reflect this limit when visualizing interaction strengths derived from the simulation.

5.2 Analytical Diffuse Interreflections

By exploiting the fixed-radius neighbors search, we can now determine all neighbors of an atom in real-time during rendering. Thus, we are able to evaluate Equation 4 within a fragment shader in the rendering stage of the visualization pipeline. This can in principle be combined with any atom rendering approach, but we have chosen to integrate it with ray-casted instancing introduced by Falk et al. [13].

To realize the interreflection effects using analytic expressions, we exploit a deferred shading approach. In the first pass, we generate a depth map and an atom ID for the current molecule and the current time frame. Thus, we can limit the evaluation of \hat{L}'' in the second pass to only those fragments which are visible from the current view-point. In the second pass we then use the generated depth map to compute the world space position for all pixels and finally evaluate the analytic expression approximating \hat{L}'' . Therefore, we use two data buffers containing atom geometry, for $c_0 + r_0$ as well as atom reflection properties f_0 , which are based on the CPK coloring [10]. The recomputed world space position is used to lookup the neighbors in our fixed-radius neighbors search data structure, whereby for each neighbor we also lookup the geometry $c_1 + r_1$ as well as the reflection properties f_1 . Based on the center of the current atom, c_0 , and the center of the neighboring atom, c_1 , we can compute d and derive d' . Thus, we have all parameters required to evaluate Equation 4, and to compute the indirect illumination affecting the current pixel. To be able to exaggerate the interreflection effects, we further apply an interreflection exponent e , which has been chosen to be 1.5 for the results in this paper, unless otherwise stated.

Since the interreflection function has been derived using only two spheres, a mechanism is required to combine the interreflections of several atoms. When dealing with one neighbor a_1 only, the proper way of calculating the resulting radiance of the fragment is to multiply

the interreflection with the environmental radiance to scale the nominal value accordingly. For multiple atoms we follow the same principle, and compute the total interreflection as $IR_{tot} = \prod IR_i^e$. Thanks to each contributions normalized format, we have found that straightforward multiplication works best when combining the individual contributions from the atoms in the vicinity.

So far, we have only taken into account the information derived from the data provided by our path tracer (cf. Section 4.1). However, as the path tracer has been simplified to operate on two atoms only, so far occlusions between neighboring atoms are not considered during rendering time, which can potentially result in overbleeding. When using a cell size of 2\AA for the fixed-radius neighbors search, the lack of occlusion does not influence the visual results. Since hydrogen with 1.2\AA is the smallest possible vdW sphere, 2.4\AA would be the minimum distance between an atom surface and the surface of an occluded neighbor, and thus the occlusion is resolved through the fixed-radius neighbors search. While even for larger cell sizes, we rarely experienced overbleeding effects, we have realized an occlusion weighting, which can be used to resolve potential effects. The technique resolves occlusions of neighboring atoms by approximating their degree of overlap. To do so, for each of the 27 cells in the vicinity of the current atom a_0 , we compute for each a_i the occlusion with all its neighbors a_j . Therefore, we sort the a_j based on the distance to a_0 , and compute a weighting factor $dp_{i,j} = \max(0, |c_0 - c_i| \cdot |c_0 - c_j|)$. $dp_{i,j}$ is computed for a_i by taking into account each a_j in the same cell. Based on all these $dp_{i,j}$ for a given atom a_i we can compute the weighting factor $w_i = \prod_{a_j} 1.0 - 0.5 \cdot dp_{i,j}$ that can then be used to weight the influence of a_i on a_0 . As this is an operation quadratic with respect to the number of atoms per cell, we analyze the performance impact of this approach in Section 6.2.

The obtained interreflection can then be combined with the reflectance f_0 of the current atom a_0 by exploiting any of the standard shading models. The pictures within this paper have been generated by applying Fresnel shading, in order to obtain subtle specular highlights. Furthermore, while the proposed model is accurate on an atomic-scale, it can be combined with more large-scale models to better capture global effects.

5.3 Ambient Occlusion

In order to emphasize atom-atom occlusions, we further propose an object-space AO approach. We exploit the fact that all geometry in the scene are spheres. Thus, we can compute the solid angle ω of the fraction of sphere S_1 as seen from x lying on sphere S_0 analytically. The required calculations are illustrated in Figure 7. While the entire solid angle Ω of S_1 from x can be written as $\Omega = 2\pi(1 - \cos(\theta))$, we only need a fraction of this angle. Therefore, we compute $t = \vec{n} \cdot \vec{v}$, whereby $\vec{v} = x - c_1$. To determine the fraction of S_1 lying above the horizon, we clamp the sum of t and r_1 to $[0, 2 \cdot r_1]$ before we normalize it. This normalized value s can then be used to compute the AO contribution as $AO = 1.0 - s \cdot (1.0 - \cos(\theta))$, which represents the amount of the hemisphere that is not occluded. The method is an approximation to the true AO as it does not take into account the cosine weight and the solid angle does not fall off in a linear fashion when the sphere is grazing the

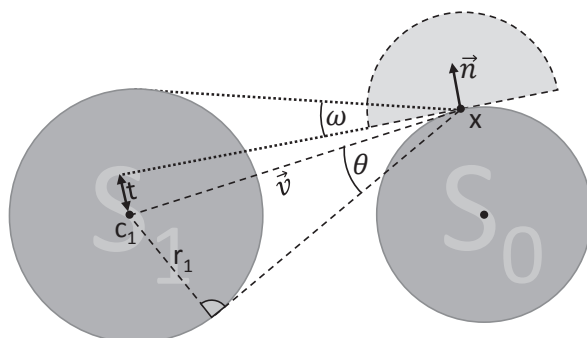


Fig. 7. The proposed ambient occlusion technique exploits the solid angle ω of the visible parts of the neighboring sphere S_1 , and relates it to the hemisphere covering the current position x .

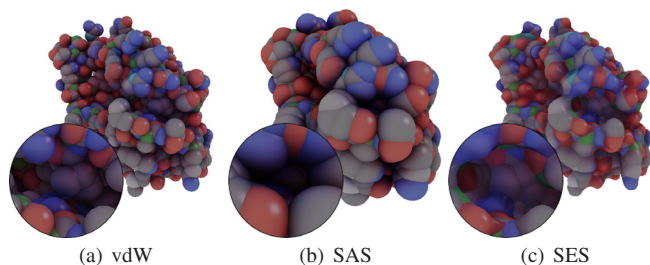


Fig. 8. Application of the proposed technique to the 3RH8 molecule for (a) van der Waals (vdW), (b) solvent-accessible surface (SAS) and (c) solvent-excluded surface (SES) representations. While vdW and SAS are directly supported, SES requires an additional interpolation.

horizon.

The strengths of the function however are that it can be evaluated anywhere in space and is not defined on the surface of a sphere, and that its precision is only limited by the resolution of the screen as it is computed per fragment. These properties give the AO a smooth falloff and crisp details, without sampling artifacts that can occur when using occupancy-based methods that are bound to the resolution of the grid. Furthermore, the proposed technique enables a direct application to SES. Just like the interreflection function, the method is dependent on the position and radii of the neighboring spheres, which makes it a local function that only takes in to account the spheres within a certain distance from the evaluated fragment. However, this method can also easily be combined with a global scale AO approaches, such as occupancy based methods to enhance small scale shadow details.

To combine the AO values with the interreflection, we first have to combine the AO of several neighbors. This is achieved through multiplication of the individual AO values as $AO_{tot} = \prod AO_i^b$, where b is an AO exponent similar to e . To later on combine the AO contribution with the interreflections, we multiply them with the incoming environmental light: $L = IR_{tot}AO_{tot}Li$.

5.4 SES Extension

The visualization algorithm presented in Section 5.2 can be directly applied to SAS and vdW representations. However, in contrary to the AO contribution, to apply it to SES slight modifications are required. Therefore, we would in the first pass render the SES mesh, whereby we write out for each fragment the ID of the closest atom as well as the surface normal of the SES mesh. In the second pass, we can then proceed as described above to obtain interreflections on SES meshes. For the parts of the SES mesh, which do not coincide with a spherical geometry including an atom, the contributions of the next three atoms can be interpolated to further improve the visual quality.

Figure 8 shows the visual results when applying our technique to SES representations as compared to vdW and SAS. All images show the 3RH8 molecule visualized with the CPK color mapping. While the proposed technique is directly applicable to vdW and SAS, SES requires the modifications described above, and can be considered an approximation only.

6 RESULTS AND LIMITATIONS

In this section, we describe the results that can be obtained when applying the proposed methods. In Section 6.1 we discuss the visual results we have generated by simulating diffuse interreflections as well as for the visualization of interaction strengths, with feedback provided by the expertise of the co-author Prof. Guallar. In Section 6.2 we analyze the performance of the presented techniques. Finally in Section 6.3, we discuss the shortcomings of the technique and how these may be addressed.

6.1 Visual Results

Diffuse interreflections. We have applied the techniques for realizing diffuse interreflections to several molecular models. Especially when analyzing molecule-ligand interactions, a good spatial comprehension is required, and the communication of atom distances can be important. Current molecular rendering packages usually provide tools to calculate

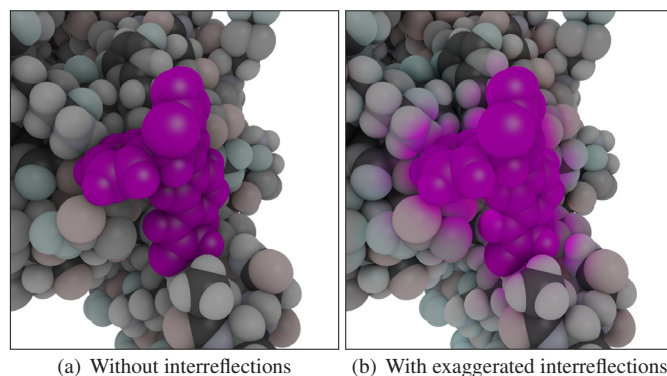


Fig. 9. A ligand binding on a receptor molecule without (a) and with exaggerated interreflections (b). In b the interreflections have been exaggerated by applying an exponent of 1.5, such that the vicinity of ligand and molecule are emphasized, and the spatial extend of the ligands upper leg becomes more apparent.

distances. However, these have to be enabled manually by the user, on a per-atom basis, which is cumbersome and time consuming. Instead diffuse interreflections generate visual cues that depict the 3D arrangement of atoms, and thus inherently communicate the closeness of atoms. Figure 1 shows three time steps from a Monte Carlo binding simulation between a Nuclear Hormone Receptor (PDB ID: 3VHV), which is depicted using desaturated CPK coloring [10], and a benzoxazin-3-one derivative ligand (PDB ID: LD1), which is depicted in magenta. As it can be seen in Figure 1 (a) it becomes clearly visible how the ligand approaches the receptor molecule. For instance, it can be seen that the three diagonally aligned receptor atoms, in the center of the figure right below the ligand, are not adjacent to the ligand, but rather covered by it. While these atoms are occluded by the ligand in Figure 1 (b), it can be seen how the ligand starts penetrating the receptor. In Figure 1 (c), it then becomes clear how the ligand submerges sideways, as some receptor atoms on the right partially occlude the ligand, and partially receive color bleeding from it.

Another molecule-ligand interaction is depicted in Figure 9, where desaturated CPK coloring is again used for the receptor. Figure 9 (a) shows the scene without interreflections, Figure 9 (b) shows the application of the presented interreflection together with an enhanced AO. The contrast communicating the 3D structure of the receptor is emphasized in Figure 9 (b), also the initial penetration of the ligand is communicated through interreflections. Furthermore, as the upper leg of the ligand also reflects on the molecule, it becomes more apparent in Figure 9 (b), that this part of the ligand is *raised*.

Since the visual quality on the atom level is one of the main benefits of the presented approach, as compared to grid sampling-based techniques, we have also applied it to single molecules consisting of fewer atoms. Figure 10 shows the application to a testosterone molecule without and with interreflections. While the effects seem to be quite subtle on the first glance, flipping these two images forth and back actually reveals that the interreflections have a rather strong impact on perceiving how the atoms are located with respect to each other,

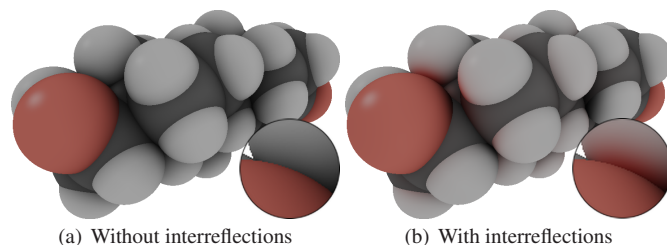


Fig. 10. Visualization of a testosterone molecule without (a) and with analytic interreflections (b). The color bleeding effects, emphasized in the inset, make the 3D structure more apparent.

especially when focusing on the oxygen atom in the front, which is also scaled up in the inset.

Encoding interaction strength. One of the major objectives of molecular simulation is the understanding of receptor-ligand binding. In molecular simulations in the context of pharmacology, for instance, docking and/or dynamical induce fit simulations are generated to analyze the binding properties of drugs. The main goal is to rationally understand the interaction energies between the drug and the protein receptor, the target, in order to propose drug modifications with better specificity and interaction strength. Interaction energies, mainly based on Coulomb, vdW and desolvation terms, appear in the form of attractive or repulsive forces, depending on atomic charges and other factors. Given a protein-ligand conformation, the spatial distribution of the atoms in both the protein and the ligand, and the combination of these energies will qualitatively determine the binding strength. Thus, obtaining better and quicker knowledge of these drug-receptor interactions directly influences the domain expert decision on chemical modifications, which can significantly improve the design of new drugs. Nowadays, the visual communication of interaction forces is mainly limited to the manual establishment of distances between atoms and molecules (residues etc.) and a subsequent calculation of the interaction strengths. By exploiting the analytic expressions derived within this paper, we can visually communicate repulsive, attractive, or the total amount of energies that are affecting certain configurations. Therefore, we exploit \hat{L}'' to encode the interaction strength on the surfaces of atoms. By exploiting the directional nature of mutual interreflections, we do not only provide means to understand the forces that are interacting but also visualize their strengths, and what atoms are contributing to these forces. Furthermore, the distance of the atoms is inherently communicated through the falloff. As interaction strengths are typically computed for a radius of 12\AA (larger distances could be chosen to capture possible long-range forces), we have adapted our cell size for the fixed-radius nearest neighbors search accordingly. Furthermore, as interaction strength computed for two atoms are independent of atoms lying in between these two atoms, we do not use any occlusion compensation as discussed in Section 5.2.

Figure 11 shows the application of the proposed interaction strength visualization to a simulation of the interaction of an aspirin ligand with the receptor molecule Phospholipase A2 (PDB ID: 1OXR). To give the ligand a canvas-like appearance, we use an exaggerated white material function for it. The interaction strengths are computed as the resulting total in each ligand atom from each protein atom. To obtain a directional component, these computations are carried out per fragment of the ligand atoms. We have encoded interaction strengths exceeding 0.5kcal in red for attractive forces, i.e., negative values, and blue for attractive forces, i.e., positive values. To communicate the absolute strength of the interaction, we exploit the interreflection exponent e (cf. Section 5.2). Thus, the ligand atoms exhibit blue, red or mixed spots indicating the direction and the absolute interaction strength. It can for instance be seen in Figure 11 (a), that an attractive force is affecting the ligand. Furthermore, in Figure 11 (b) it can be seen that both, repulsive and attractive forces, are influencing one atom of the ligand.

Figure 12 shows two time steps of the Nuclear Hormone Receptor simulation. The purpose of this simulation is to investigate the effects of the receptor molecule dynamics on the ligand binding mechanism.

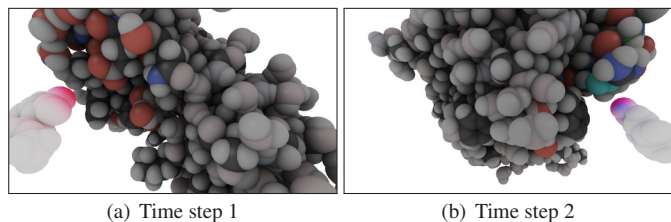


Fig. 11. Visualization of the interaction strength between a ligand and a receptor molecule for two time steps of an aspirin binding simulation. Attracting forces between atoms are depicted in red, while repulsive forces are depicted in blue. In contrast to existing techniques, our approach enables us to visualize all atom/atom intermolecular interactions in a directional manner, without occluding the scene with additional geometry, such as lines or rulers.

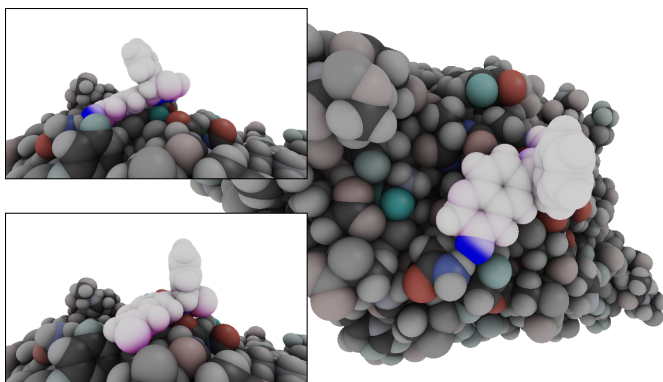


Fig. 12. Visualization of the interaction strength around a Nuclear Hormone Receptor. The insets show the ligand from a different perspective, whereby the upper inset shows the same time step as shown in the overview, and the lower inset shows the next time step. The red/blue coding enables qualitative prediction of the next steps in the simulation, and thus inherently captures the dynamics of the ligand.

Since the strong repelling forces of a hydrogen-hydrogen close contact dominate in Figure 12 (top inset), it can be expected that the ligand will be displaced towards the right within the next time step of the simulation. Indeed, in Figure 12 (bottom inset) (the next time step), a large displacement can be observed into the predicted direction. Furthermore, now the attractive forces seem to be predominant, which results in a *sliding* of the ligand along the molecule's surface. In fact that is what we could observe in the subsequent frames. To our knowledge, these are the first visualizations capable of capturing these simulation dynamics in a static frame, and thus enabling interesting insights into the simulation. Current techniques only enable the depiction of distances and charge values through rulers or lines, which results in a cluttering of the scene.

When analyzing these visualizations shown in Figure 11 and Figure 12, it is further important to know which atoms of the receptor molecule are interacting with the ligand in the current time frame. Therefore, we do not visualize the interaction strength on the receptor,

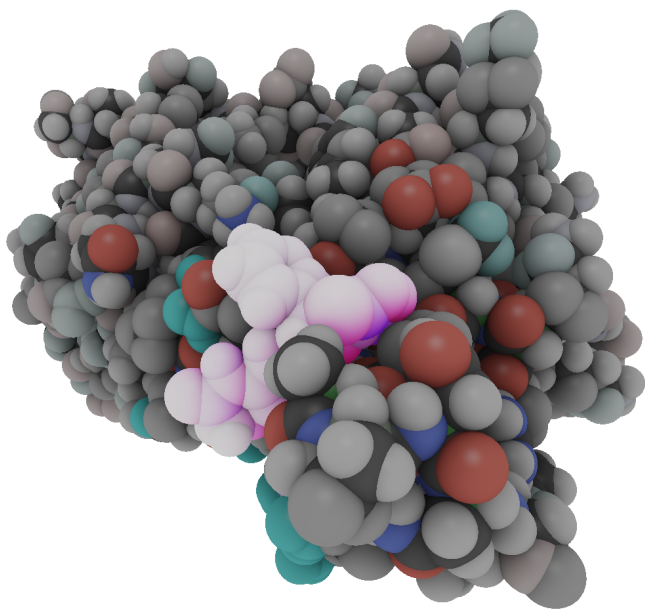


Fig. 13. Visualization of the interaction strength between a ligand and a receptor molecule, whereby the atom identity color of the affecting receptor atom is used to enhance ligand atoms. Thus, it becomes possible to determine, which receptor atoms have the largest impact on the ligand.

Molecule	PDB ID	Atom Count	2Å		6Å		12Å
			Off	On	Off	On	Off
Testosterone	TES	49	514	474	244	133	249
Isomerase	IOGZ	1057	427	413	112	74	58
Photoreceptor	3RH8	2526	351	317	81	58	43
Maltoporin	1AF6	10517	298	291	104	72	46

Table 1. Frame rate measurements (FPS) for different molecules obtained at an image resolution of 1280×720 pixels with varying cell size (2Å, 6Å, and 12Å). For 2Å and 6Å, the two columns compare the performance without (Off) and with [On] occlusion compensation.

but rather use the standard CPK color mapping to support identification of atom types. For those receptor atoms not interacting with the ligand in the current time step, desaturation is applied such that their types can still be identified, but the focus is on those atoms responsible for the interaction.

As an alternative coloring for the ligand, we propose to use the actual CPK color mapping of the interacting receptor atoms. Thus, it becomes clear which receptor atoms have the greatest contribution to the interaction. Figure 13 shows a visualization generated with this technique. As it can be seen, now the origin of the acting force can be deciphered, as the visualization uses the atom type coloring. To quantify the absolute interaction strength, it can also be used to proportionally saturate the resulting bleeding effects.

So far, in the visualizations discussed in this subsection, we have only incorporated Coulomb and vdW forces, and we leave the addition of Solvent energies for future work.

Expert feedback. As interatomic forces, both from a classical or quantum perspective, depend on the distance between particles, the techniques presented here constitute an excellent choice to visualize such interactions. The presented interaction strength encoding (see Figure 11) indicates the potential in exploring the attractive and repulsive forces on a Monte Carlo ligand diffusion simulation. By doing so, specialized researchers working on drug design can quickly and intuitively understand which parts of the drug are more relevant and which others are better candidates for chemical modifications. Accordingly, the encoding can also serve as a natural vehicle for the following parameters: i) molecular bleeding (by summing up atomic contributions) for a quick estimate of the overall ligand binding energy; ii) individual non-bonding force field terms bleeding, to better detect long-range forces and adjust force field parameters; iii) electronic polarization bleeding derived from quantum methods or polarizable force fields, to visualize quantum effects in molecular association.

6.2 Performance Results

To analyze the runtime of the presented approach, we have conducted a performance analysis on an Intel Core i5, 3.2GHz Quad-Core System with 8 GB of RAM, and an Nvidia GeForce GTX960 graphics card equipped with 2GB of RAM. As shown in Table 1, we have measured frame rates for several molecules varying in complexity with respect to their atom count. Furthermore, we have varied the cell size used for the fixed-radius neighbors search, whereby we have tested 2Å, 6Å and 12Å, whereby the latter is the radius typically used in molecular dynamics simulations. For all cell sizes, we have measured the frame rates at an image resolution of 1280×720 pixels. For the cell sizes of 2Å and 6Å, we have further switched on and off the occlusion compensation as discussed in Section 5.2. For the 12Å cell size, we have not performed these measurements, as this radius is only relevant for depicting the interaction strength, which requires the occlusion compensation to be always off. For all measurements we have chosen a zoom level such that the respective molecule fills an as large part of the viewport as possible without being clipped by the borders.

As it can be seen in Table 1, we achieve interactive frame rates in all cases. As expected, the occlusion compensation results in a frame rate penalty, which is proportional to the atom count and the cell size. In fact, for a cell size of 12Å the frame rates for the more complex molecules were below 10 fps. Nevertheless, since 12Å are only used to illustrate atom/atom interaction strengths, it is mandatory to switch off the occlusion compensation in these cases anyway. The table also

shows, that the frame rates also dependent on the atom count, whereas this is not the main influencing parameter.

6.3 Limitations

In this section we briefly discuss the main limitations of our approach.

Cell size and neighborhood. If the cell size is too small, few atoms will be found through the neighboring cells which will degrade the visual quality of the technique. On the contrary, if the cell size is too big, too many atoms will have to be processed per fragment and the performance will suffer. In general, we have found that a cell size of 3-6Å seem to provide a good trade off between visual quality and performance when using the technique for rendering.

Overbleeding. As mentioned in section 6.2 overbleeding can occur because of the techniques being applied in a pair-wise frame and does not take into account third party atoms that might be occluding. This is can be compensated to some extent with the use of occlusion compensation as mentioned in Section 5.2. However, in most cases this is not necessary, with the exception of highly exaggerated interreflection strengths.

AO vs. diffuse interreflections. Currently we use both AO and interreflections when computing the resulting picture. In a perfect scenario with an expression \hat{L}'' that fits perfectly to the path-traced data we would not need the AO model. However the fit of the expression we use is not perfect and will probably never be. It generally lacks some of the falloff that occurs when two atoms are close and occlude each other more than is added through interreflection. This scenario can be seen in the left most example in Figure 4, and the two darkest curves in Figure 5 (a). Therefore we combine the AO and the diffuse interreflections as mentioned in Section 5.3.

7 CONCLUSIONS AND FUTURE WORK

In this paper we have proposed a novel visualization approach for space filling molecular representations. Inspired by the physical processes underlying light transport, and motivated by the perceptual benefits of interreflections, the presented visualization techniques capture the diffuse interreflections between atoms of the same or different molecules. While the computational complexity of these effects would forbid to apply them during interactive visualization, our approach exploits symbolic regression to find an analytical solution to this problem. By reducing the problem to a simplified two atom setup, we are able to derive a five parameter function, which captures these light interactions adequately. Symbolic regression then enabled us to find an analytic expression, which can be implemented on modern GPUs in order to achieve interactive frame rates. We have derived the underlying theory, discussed the interactive visualization algorithm, and presented results showing the benefits of the proposed technique. Our results indicate, that the proposed approach can be used for both, improving the degree of realism of molecular visualizations, and for visually communicating relevant simulation parameters.

In the future we would like to investigate several aspects which are beyond the scope of this paper. First, we would like to employ a better SES interpolation, which takes into account the curvature along the surface rather than just the vicinity to atoms, and investigate how to extend to other space filling representations, such as ligand-excluded surfaces (LES) [33] or molecular skin surface (MSS). An obvious endeavor is also the extension to other primitive types, such as cylinders for visualizing bonds. Furthermore, we would like to combine the proposed approach with semi-transparency, as this would enable us to give an occlusion-free view to influencing atoms which sit deep inside the receptor molecule. To provide an even better visualization with respect to possible ligand modifications, we would like to experiment with techniques for combining positive and negative interaction forces together with influencing atom identities in the same visualization. Finally, we aim to investigate which other areas of visualization that are based on multiple parameters that can benefit from symbolic regression analysis.

ACKNOWLEDGMENTS

All static molecules have been obtained from the RCSB Protein Data Bank (<http://www.rcsb.org>), while the dynamic binding studies have been computed using the PELE Monte Carlo solver (<http://pele.bsc.es>) in cooperation with the pharmaceutical company AstraZeneca. This research is partially supported by TIN2014-52211-C2-1-R from the Spanish Government.

REFERENCES

- [1] J. Behley, V. Steinhage, and A. B. Cremers. Efficient radius neighbor search in three-dimensional point clouds. In *IEEE International Conference on Robotics and Automation (ICRA)*, 2015. to appear.
- [2] M. G. Bloj, D. Kersten, and A. C. Hurlbert. Perception of three-dimensional shape influences colour perception through mutual illumination. *Nature*, 402(6764):877–879, 1999.
- [3] A. Bryden, G. Phillips Jr., and M. Gleicher. Automated illustration of molecular flexibility. *IEEE Transactions on Visualization and Computer Graphics*, 18(1):132–145, Jan. 2012.
- [4] M. Bunnell. Dynamic ambient occlusion and indirect lighting. *Gpu gems*, 2(2):223–233, 2005.
- [5] M. K. Chandraker, F. Kahl, and D. J. Kriegman. Reflections on the generalized bas-relief ambiguity. In *IEEE Computer Vision and Pattern Recognition (CVPR)*, volume 1, pages 788–795, 2005.
- [6] M. Chavent, B. Lévy, M. Krone, K. Bidmon, J.-P. Nominé, T. Ertl, and M. Baaden. GPU-powered tools boost molecular visualization. *Briefings in Bioinformatics*, 12(6):689–701, 2011.
- [7] G. Cipriano and M. Gleicher. Molecular surface abstraction. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1608–1615, Nov. 2007.
- [8] G. Cipriano, G. N. Phillips Jr., and M. Gleicher. Multi-scale surface descriptors. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):1201–1208, Nov. 2009.
- [9] E. Clementi and D.-L. Raimondi. Atomic screening constants from SCF functions. *The Journal of Chemical Physics*, 38(11):2686–2689, 1963.
- [10] R. B. Corey and L. Pauling. Molecular models of amino acids, peptides, and proteins. *Review of Scientific Instruments*, 24(8):621–627, 1953.
- [11] O. Daee Lampe, I. Viola, N. Reuter, and H. Hauser. Two-level approach to efficient visualization of protein dynamics. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1616–1623, Nov. 2007.
- [12] M. Falk, M. Klann, M. Reuss, and T. Ertl. Visualization of signal transduction processes in the crowded environment of the cell. In *Proceedings of the 2009 IEEE Pacific Visualization Symposium*, pages 169–176, Washington, DC, USA, 2009. IEEE Computer Society.
- [13] M. Falk, M. Krone, and T. Ertl. Atomistic visualization of mesoscopic whole-cell simulations using ray-casted instancing. *Computer Graphics Forum*, pages 195–206, 2013.
- [14] T. D. Goddard and T. E. Ferrin. Visualization software for molecular assemblies. *Current Opinion in Structural Biology*, 17(5):587–595, 2007.
- [15] S. Grottel, P. Beck, C. Müller, G. Reina, J. Roth, H.-R. Trebin, and T. Ertl. Visualization of electrostatic dipoles in molecular dynamics of metal oxides. *IEEE Transactions on Visualization and Computer Graphics*, 18(12):2061–2068, Dec. 2012.
- [16] S. Grottel, M. Krone, C. Müller, G. Reina, and T. Ertl. MegaMol—a prototyping framework for particle-based visualization. *IEEE Transactions on Visualization and Computer Graphics*, 21(2):201–214, 2015.
- [17] S. Grottel, M. Krone, K. Scharnowski, and T. Ertl. Object-space ambient occlusion for molecular dynamics. In *IEEE Pacific Visualization Symposium*, pages 209–216, 2012.
- [18] S. Grottel, G. Reina, C. Dachsbaecher, and T. Ertl. Coherent Culling and Shading for Large Molecular Dynamics Visualization. *Computer Graphics Forum*, 29(3):953–962, 2010.
- [19] M. D. Hanwell, D. E. Curtis, D. C. Lonie, T. Vandermeersch, E. Zurek, and G. R. Hutchison. Avogadro: An advanced semantic chemical editor, visualization, and analysis platform. *J. Cheminformatics*, 4:17, 2012.
- [20] R. C. Hoetzlein. Fast fixed-radius nearest neighbors: Interactive million-particle fluids. In *GPU Technology Conference*, 2014.
- [21] W. Humphrey, A. Dalke, and K. Schulten. VMD: visual molecular dynamics. *Journal of Molecular Graphics*, 14(1):33–38, 1996.
- [22] J. T. Kajiya. The rendering equation. In *Proceedings of the 13th Annual Conference on Computer Graphics and Interactive Techniques, SIGGRAPH '86*, pages 143–150, New York, NY, USA, 1986. ACM.
- [23] N. A. Khazanov and H. A. Carlson. Exploring the composition of protein-ligand binding sites on a large scale. *PLoS Computational Biology*, 9(11):e1003321, 2013.
- [24] A. Knoll, I. Wald, P. A. Navrátil, M. E. Papka, and K. P. Gauthier. Ray tracing and volume rendering large molecular data on multi-core and many-core architectures. In *Proceedings of the 8th International Workshop on Ultrascale Visualization, UltraVis '13*, pages 5:1–5:8, New York, NY, USA, 2013. ACM.
- [25] S. Kottavel, M. Falk, E. Sundén, and T. Ropinski. Coverage-Based Opacity Estimation for Interactive Depth of Field in Molecular Visualization. In *Proceedings of the 2015 IEEE Pacific Visualization Symposium*, pages 255–262, 2015. to appear.
- [26] B. Kozlikova, M. Krone, N. Lindow, M. Falk, M. Baaden, D. Baum, I. Viola, J. Parulek, and H.-C. Hege. Visualization of Biomolecular Structures: State of the Art. In R. Borgo, F. Ganovelli, and I. Viola, editors, *Eurographics Conference on Visualization (EuroVis) - STARs*. The Eurographics Association, 2015.
- [27] M. Krone, K. Bidmon, and T. Ertl. Interactive visualization of molecular surface dynamics. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):1391–1398, Nov. 2009.
- [28] M. Langer. Interreflections. In K. Ikeuchi, editor, *Computer Vision*, pages 423–424. Springer US, 2014.
- [29] K. Lawonn, M. Krone, T. Ertl, and B. Preim. Line Integral Convolution for Real-Time Illustration of Molecular Surface Shape and Salient Regions. *Computer Graphics Forum*, 33(3):181–190, 2014.
- [30] M. Le Muzic, J. Parulek, A. K. Stavrum, and I. Viola. Illustrative visualization of molecular reactions using omniscient intelligence and passive agents. *Computer Graphics Forum*, 33(3):141–150, 2014.
- [31] N. Lindow. Dynamic channels in biomolecular systems: Path analysis and visualization. In *Proceedings of the 2012 IEEE Symposium on Biological Data Visualization (BioVis)*, BIOVIS '12, pages 99–106, Washington, DC, USA, 2012. IEEE Computer Society.
- [32] N. Lindow, D. Baum, and H.-C. Hege. Interactive Rendering of Materials and Biological Structures on Atomic and Nanoscopic Scale. *Computer Graphics Forum*, 31(3pt4):1325–1334, 2012.
- [33] N. Lindow, D. Baum, and H.-C. Hege. Ligand excluded surface: A new type of molecular surface. *IEEE Transactions on Visualization and Computer Graphics*, 20(12):2486–2495, 2014.
- [34] T. Luft, C. Colditz, and O. Deussen. Image enhancement by unsharp masking the depth buffer. *ACM Transactions on Graphics*, 25(3):1206–1213, July 2006.
- [35] K. Momma and F. Izumi. VESTA: a three-dimensional visualization system for electronic and structural analysis. *Journal of Applied Crystallography*, 41(3):653–658, 2008.
- [36] K. Momma and F. Izumi. VESTA 3 for three-dimensional visualization of crystal, volumetric and morphology data. *Journal of Applied Crystallography*, 44(6):1272–1276, 2011.
- [37] J. Parulek, D. Jönsson, T. Ropinski, S. Bruckner, A. Ynnerman, and I. Viola. Continuous Levels-of-Detail and Visual Abstraction for Seamless Molecular Visualization. *Computer Graphics Forum*, 33(6):276–287, 2012.
- [38] R. Ramamoorthi and P. Hanrahan. An efficient representation for irradiance environment maps. In *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, pages 497–500. ACM, 2001.
- [39] K. Reda, A. Febretti, A. Knoll, J. Aurisano, J. Leigh, A. Johnson, M. Papka, and M. Hereld. Visualizing large, heterogeneous data in hybrid-reality environments. *IEEE Computer Graphics & Applications*, 33(4):38–48, July 2013.
- [40] G. Reina and T. Ertl. Hardware-accelerated glyphs for mono- and dipoles in molecular dynamics visualization. In *Proceedings of the Seventh Joint Eurographics / IEEE VGTC Conference on Visualization, EUROVIS'05*, pages 177–182, Aire-la-Ville, Switzerland, Switzerland, 2005. Eurographics Association.
- [41] P. Ren, J. Wang, M. Gong, S. Lin, X. Tong, and B. Guo. Global illumination with radiance regression functions. *ACM Transactions on Graphics*, 32(4):130:1–130:12, July 2013.
- [42] A. Sarikaya, D. Albers, J. Mitchell, and M. Gleicher. Visualizing validation of protein surface classifiers. *Computer Graphics Forum*, 33(3):171–180, 2014.
- [43] K. Scharnowski, M. Krone, G. Reina, T. Kulschewski, J. Pleiss, and T. Ertl. Comparative visualization of molecular surfaces using deformable models. *Computer Graphics Forum*, 33(3):191–200, 2014.
- [44] M. Schmidt and H. Lipson. Distilling free-form natural laws from experimental data. *Science*, 324(5923):81–85, 2009.
- [45] M. Tarini, P. Cignoni, and C. Montani. Ambient occlusion and edge cueing for enhancing real time molecular visualization. *IEEE Transactions on Visualization and Computer Graphics*, 12(5):1237–1244, 2006.
- [46] M. van der Zwan, W. Lueks, H. Bekker, and T. Isenberg. Illustrative molecular visualization with continuous abstraction. *Computer Graphics Forum*, 30(3):683–690, 2011.
- [47] J. Weber. ProteinShader: illustrative rendering of macromolecules. *BMC Structural Biology*, 9(1), 2009.
- [48] I. Yu, A. Cox, M. H. Kim, T. Ritschel, T. Grosch, C. Dachsbaecher, and J. Kautz. Perceptual influence of approximate visibility in indirect illumination. *ACM Transactions on Applied Perception*, 6(4):24:1–24:14, Oct. 2009.